Survey of Decentralized Control Methods for Large Scale Systems

NILS R. SANDELL, JR., MEMBER, IEEE, PRAVIN VARAIYA, MEMBER, IEEE, MICHAEL ATHANS, FELLOW, IEEE, AND MICHAEL G. SAFONOV, MEMBER, IEEE

Abstract—**This paper surveys the control theoretic literature on decentralized and hierarchical control, and methods of analysis of large scale systems.**

I. Introduction to the Survey Paper

OVER THE past twenty-five years engineers have invented a variety of procedures for analyzing systems and for designing control strategies. These procedures may be classified into three types.

1) Procedures for modeling dynamical systems (state space formulation, input-output transfer function descriptions, etc.)

2) Procedures for describing qualitative properties of system behavior (controllability, stabililty, observability, etc.)

3) Procedures for controlling system behavior (stabilizing feedback, optimal control, etc.).

All of these procedures rest on the common presupposition of *centrality*; all the information available about the system, and the calculations based upon this information, are centralized, that is, take place at a single location. It is useful to distinguish two kinds of available information:

1) Information about the system model; (we may call this "off-line" or *a priori* information)

2) Sensor information about the system response, that is, at each time the set of all measurements on the system made up to that time.

We stress that from a theoretical point of view the notion of centrality is common to both the classical servomechanism approach and the modern estimation and control theory approach. Whether or not one designs a simple lead-lag network using a Nichols chart, or a stochastic regulator using dynamic programming one uses theoretical tools that presuppose centrality.

Manuscript received June 6, 1977; revised October 27, 1977. This work was supported by ERDA under Grant ERDA-E (49-18)-2087, by ONR Contract N00014-76-C-0346, and by Joint Services Electronics Program F44620-76-C-0100.

N. R. Sandell, Jr., and M. Athans are with the Electronic Systems Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139.

P. Varaiya is with the Electronics Research Laboratory, University of California, Berkeley, CA 94720.

M. G. Safonov was with the Electronic Systems Laboratory, Massachussetts Institute of Technology, Cambridge, MA 02139. He is now with the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90007.

When considering large systems the presupposition of centrality fails to hold due either to the lack of centralized information or the lack of centralized computing capability. There are many examples of large scale systems that present a great challenge to both system analysts and control system designers. As a partial list of such problems we mention

a) power networks
b) urban traffic networks
c) digital communications networks
d) flexible manufacturing networks
e) ecological systems
f) economic systems

as typical examples. Such physical systems are often characterized by geographical separation so that issues such as the economic cost and reliability of communication links have to be taken into account, thus providing impetus for a decentralized scheme. The technological advances and reduced cost of microprocessors provides another technological force for distributed computation. In both civilian and military command-and-control systems distributed data bases (with or without duplication for reliability) present new challenges for decentralized decision making.

Thus, for economic and possibly reliability reasons, there is a trend for decentralized decision making, distributed computation, and hierarchical control. However, these desirable goals of structuring a distributed information and decision framework for large scale systems do not "mesh" with the available centralized methodologies and procedures associated with classical and modern control theory, and the available tools for design cannot be directly applied. This paper is a survey of the techniques which have been developed in the control theoretic literature (or can be regarded as such) in response to the failure of centrality. The survey naturally divides into three parts. In Section II we look at techniques aimed at simplifying model descriptions; in Section III we discuss the procedures for testing stability; and in Sections IV and V we study control techniques with decentralized control problems discussed in the former and hierarchical control problems in the latter.

A survey such as this is bound to reflect the ignorance and biases of its authors. This is particularly a problem here, since there is no precise definition of what exactly constitutes large scale system theory. We apologize for overlooking papers which should have been included as well as for any misleading emphasis.

0018-9286/78/0200-0108$00.75

II. Model Simplification

A. Introduction

It is a common procedure in engineering practice to work with mathematical models that are simpler, but less accurate, than the best available model of a given physical process. There are two quite distinct motivations for this practice. The first is to reduce the computational burden associated with simulation, analysis and control system design. The second is based on the realization that a simplified model will lead to a simplified control system structure. To see that these indeed represent two distinct issues, consider for example the problem of reduced order Kalman filter design [1]. In order to determine the error covariance matrix for a Kalman filter based on a reduced order model, it is necessary to solve covariance equations of order $(n+n_r)^2$, where n is the number of states for the full order model and n_r is the number of states for the reduced model. Since the covariance of an optimal Kalman filter is determined by covariance equations of order n^2, reduced order models would be of very little interest in Kalman filter design if reduced *off-line* computation were the only issue. Of course, a filter based on the reduced order model has a simpler structure (reduced order) that permits a reduction in *on-line* computation in terms of the number of real-time arithmetic operations. Thus, we see that *reduced computation and simplified structure are separate desirable goals that are not necessarily compatible.*

Reduction of computation and simplification of structure are of particular concern in decentralized control of large scale systems, but are also of concern in almost all areas of control theory and its applications. Thus, our criterion for what material to survey is even more subjective here than in other sections.

The work to be surveyed is divided into two classes. The first class, called *aggregation* methods, assumes that a mathematical model of a given system can be simplified by introducing a coarser state space description which still retains some key qualitative properties of the system. The second class, called *perturbation* methods, consists of procedures in which certain dynamic interactions in the system are ignored. These two classes of model simplification methods are related as we shall see. In both cases, the underlying philosophy of the methods is outlined, and the implications of the methods for decentralized control of large scale systems is examined.

B. Aggregation Methods

The intuitive idea behind the notion of aggregation is quite simple. Suppose that S_1 is a mathematical description of a physical system using a given set of variables, and S_2 is a consistent description of the same system using a smaller set of variables. Then S_2 is termed an aggregate model for S_1, and the variables of the system S_2 are termed aggregate variables.

Interestingly, notions equivalent to that of aggregation have appeared apparently independently in at least three areas of system theory. In automata theory, the intuitive notion above is made precise by the statement that S_2 and S_1 are related by an automata homomorphism [3]. In the theory of Markovian decision processes, S_2 is said to result from S_1 by merging states of S_1 [4]. Third, the phrase "S_2 is an aggregated model of S_1" has been introduced in the control theoretic literature and made precise by Aoki [2] for linear systems.[1] There is a particularly close connection between the idea of merging and of aggregation [6]. However, we will restrict our attention to the linear system case.

Suppose then that the unaggregated system S_1 is described by the state equation

$$\dot{x}(t)=Ax(t)+Bu(t) \tag{2.1}$$

and the aggregated system S_2 is described by the state equation

$$\dot{z}(t)=Fz(t)+Gu(t). \tag{2.2}$$

In order for S_2 to be an aggregated model of S_1, we require that

$$z(t)=Cx(t) \tag{2.3}$$

for all t. This requirement is termed *dynamic exactness*. It is easy to see that dynamic exactness is achieved if and only if the matrix equations

$$FC=CA \tag{2.4}$$

$$G=CB \tag{2.5}$$

are satisfied. Analysis of these conditions reveals that dynamic exactness is achieved only when the aggregate state vector $z(t)$ is a linear combination of certain of the modes of $x(t)$. In this case, the eigenvalues of F are the eigenvalues of A corresponding to those modes of $x(t)$ which are retained in $z(t)$. Thus, the notion of aggregation for linear systems is, in fact, a generalization of the familiar idea of simplifying linear systems by retaining the dominant modes. This restricts the class of matrices C that are aggregation matrices.

Further insight into the nature of the class of matrices for which dynamic exactness can be achieved is obtained by realizing that the aggregation problem as posed for linear systems is in fact a problem of minimal realization. Notice that (2.1) and (2.3) determine a dynamic relationship between $u(t)$ and $z(t)$ that can be described by the transfer function matrix

$$H(s)=C(sI-A)^{-1}B. \tag{2.6}$$

In order for the dynamic exactness condition to hold, we must also have

$$H(s)=(sI-F)^{-1}G. \tag{2.7}$$

That is, the transfer function matrix must be realizable by

[1] There is a long history of study of the aggregation concept in the economics literature; an excellent recent survey is given by Chipman [5].

either (2.1) and (2.3) or by (2.2). Hence, if $z(t)$ has lower dimension than $x(t)$, the state description defined by (2.1) and (2.3) is nonminimal. But this is possible if and only if there are pole-zero cancellations. Thus, the class of aggregation matrices C is restricted to those that create zeros in the input-output relationship between $u(t)$ and $z(t)$ that cancel poles of the relationship. The poles cancelled by the zeros are precisely the eigenvalues of A that are not retained in F.

Up to this point, we have emphasized the role of the aggregate model as an exact representation of the response of certain linear combinations of the state variables of S_1 to the input vector $u(t)$. However, the aggregate state variables $z_i(t)$ will *not* in general correspond exactly to physical variables. Therefore, an alternative point of view, that may be more useful in applications, is to regard $z(t)$ as an approximation to physical variables. In other words, in addition to (2.2), we desire

$$z(t) \approx y(t) \triangleq Hx(t) \tag{2.8}$$

where the matrix H picks out components or linear combinations of components of $x(t)$ that are to be approximated. (This amounts to a modal approximation of $y(t)$.) Of course, the choice of the aggregation matrix C can greatly influence the nature of the approximation (2.8).

Although dynamic exactness restricts the class of possible aggregation matrices for a given system of the form (2.1), there are still many ways to choose the matrix C. Various model reduction techniques proposed by several authors can be viewed as aggregation methods with a particular choice of C implied. For example, a natural approach [7] is to require that a given mode of S_1 that is retained in S_2 be represented in $z(t)$ in the same proportions that it is represented in $y(t)$, i.e., that the mode shapes be preserved. Another criterion that has been discussed by several authors [8], [9] is that $z(t)$ should have the same steady state response to step inputs as $y(t)$. Yet another approach is to minimize some measure of the difference $z(t)-y(t)$ for a given class of inputs [10]–[12]. Any of these techniques may be of interest depending upon the intended use of the aggregate model.

One difficulty, for some potential applications, with the approaches mentioned so far is that even if the aggregate variables are given a physical interpretation from (2.8), there may be nonphysical coupling coefficients in the F matrix. For example, one study [13] of model simplification of a dynamical model of a power system, resulted in nonphysical direct couplings between the voltage regulators and governors of the machines of the system (in addition to the physical couplings through the transmission system). This suggests that in addition to the condition of dynamical exactness, some condition of structure preservation should be added to the aggregation notion. This idea has appeared independently in an automata theory context [3] and in linear systems theory [13]. Of course, a simplified model can only preserve a portion of the characteristics of a more complex model, and structure preservation can only be achieved at the expense of other properties one might wish to preserve. For example, it is, in general, not possible to preserve mode shapes for a structurally constrained aggregation [13].

As mentioned in the introduction to this section, aggregation techniques are of interest in large scale system analysis and design to the extent that valid analyses and designs can be based on the aggregate models. We will distinguish between two cases. In the first, the aggregate state $z(t)$ is of direct interest, while in the second the aggregate state is of interest only to the extent that it approximates some set of physical variables.

If the aggregate state $z(t)$ is of direct interest, it can be computed exactly from (2.2), so that simulations (including covariance simulations) can be carried out with reduced computational burden. Since the order of an optimal estimator or observer for $z(t)$ is equal to its dimension, a state reconstructor with a simplified, reduced order structure can be employed. Similarly, simplified feedback control structures, determined by linear-quadratic optimization or other means can be obtained. Note that if feedback is derived from $z(t)$, the modes not retained in the aggregate model are excited by the feedback loops but cannot destabilize the system since they do not affect $z(t)$.

If the aggregate state is of interest only as an approximation to some set of physical variables, then it is much harder to make definitive statements. Analyses can be carried out with less computational expense with the aggregate model, but the validity of the results depends on the accuracy of the approximation (2.8). The validity of this approximation in turn depends on the system and its inputs as well as the aggregation technique. For example, the aggregate model cannot correctly predict the response to an initial condition disturbance heavily concentrated in a mode that is omitted from the aggregate model. Thus the expected inputs (disturbance and otherwise) to the system should be taken into account in the selection of an aggregate model.

If the aggregate model is used to determine a simplified feedback structure, the situation becomes even more complex. Modes that are not represented in the aggregate model can interact with the feedback loops to destabilize the system. This can be a particular problem if the aggregate model was obtained by neglecting fast modes and the feedback loops of the control structure have a high-bandwidth or, in the case of sampled data control, the sampling rate is too low.

In the particular case of linear-quadratic state variable feedback optimal control, the effects of designing a regulator for an aggregate model and applying it to the original system have been examined. One approach that has been pursued is to derive bounds on the optimal performance in terms of the solution to an aggregate Riccati equation [2], [14]. Another interesting idea, proposed in [15], is to choose the aggregation matrix so that the performance of a regulator for the unaggregated system based on the aggregated model is as good as possible.

From a pragmatic point of view, aggregation issues cannot be dealt with in a purely abstract framework. *How*

to aggregate a system depends on the objectives. Even if one agrees on approximating an nth order system by an mth order system ($m<n$), then the best mth order aggregated model may be different for estimation and prediction purposes from that used for control purposes. The problem is even more cloudy when stochastic control is considered which involves both the estimation and the control problem.

What would be useful, in our opinion, is to attempt to develop a theory of aggregation which actively incorporates not only the order of the system dynamics but, in addition:

a) The performance index that has to be optimized, if any.

b) The actuator and sensor characteristics that are available.

c) The nature of the disturbances.

Such a theory is lacking. To make matters worse, the notion of complexity of a control system (as may be measured by the number of feedback loops, number of arithmetic operations, etc.) has to take into account the advances in LSI electronic components, whose inherent structure and reliability may make some of the traditional notions of complexity obsolete.

C. Perturbation Methods

Perturbation methods[2] are useful for dealing with a system that can be approximated by a system of simpler structure. Mathematically, the difference in response between the actual and approximating systems is modeled as a perturbation term driving the latter. Perturbation methods can be viewed as approximate aggregation techniques. Although in principle applicable to both linear and nonlinear problems, we shall for simplicity restrict attention to the linear case. We divide perturbation procedures into the two subclasses of *weak coupling* and *strong coupling* methods, although this terminology is not standard.

1) Weak Coupling

The weak coupling approach is based on the notion of a *nonsingular perturbation*, i.e., a perturbation term in the right-hand side of a differential equation. A typical formulation is to consider the dynamic system

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} A_{11} & \epsilon A_{12} \\ \epsilon A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B_1 & 0 \\ 0 & B_2 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \tag{2.9}$$

where ϵ is a small positive parameter. Clearly, for $\epsilon=0$, the system decouples into two completely independent subsystems; i.e., two approximate aggregate models are obtained, one for each subsystem. Computation associated with simulations, control system design, etc. is reduced since two independent lower dimensional problems are considered, and since computation generally grows at a greater than linear rate as the dimension increases. (This effect is enhanced for more than two subsystems.) Moreover, control systems designed for the decoupled problems provide a completely decentralized control structure for the original problem, as illustrated in Fig. 1.

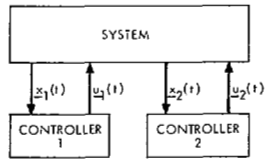

Fig. 1. Decentralized control structure.

Research in the weak coupling area has followed two lines of approach. One approach is to set $\epsilon=0$ in (2.9) and attempt to quantify the degree of approximation when, in the actual system $\epsilon\neq 0$. For example, consider a linear quadratic optimal control problem with weighting matrices block diagonal and compatible with the partitioning in (2.9). Bounds on the loss of performance incurred by implementing the decoupled solution on the actual system have been obtained in [17] and [18]. In this same work, an explicit weak coupling condition is obtained which insures that the approximate solution stabilizes the actual system. This condition is always satisfied for small ϵ, but can also be satisfied for large ϵ if the coupling is of an appropriate type. Similar results are obtained in the stochastic case in [19] and [20].

The second area of research is concerned with exploiting the weak coupling structure algorithmically. For example, consider again the linear-quadratic optimal control problem of the previous paragraph. It can be shown [21] that the Riccati equation of the problem can be expanded in powers of ϵ, and that retaining n terms in the expansion gives an approximation of order $2n$ to the optimal cost. (This result generalizes to the stochastic case as well [22].) The zeroth order terms are computed by decoupled Riccati equations obtained by setting $\epsilon=0$ in (2.9), and the higher order terms are computed by solution of decoupled linear equations. One drawback of the approach is that it is nonrecursive in character, and thus, difficult to implement. This problem can be overcome by directly exploiting the weak coupling structure to obtain an iterative algorithm; see [23] and [24]. A second drawback is that if first and higher order terms are retained in the expansion, then there is feedback from one subsystem to the other so that a decentralized solution is not obtained. Ways of dealing with this problem are discussed in the section on decentralized control.

2) Strong Coupling

The strong coupling approach is based on the notion of a *singular perturbation*, i.e., a perturbation to the left-hand side of a differential equation. Consider a dynamic system

[2]The literature on perturbation methods in applied mathematics is extensive. As in other sections of the paper, we shall concentrate on topics that have already influenced the control-theoretic literature. For a discussion of an alternate approach see [16].

of the form

$$\dot{x}_1(t)=A_{11}x_1(t)+A_{12}x_2(t) \quad \text{(slow system)} \quad (2.10)$$

$$\epsilon\dot{x}_2(t)=A_{21}x_1(t)+A_{22}x_2(t) \quad \text{(fast system)} \quad (2.11)$$

where ϵ is a small positive parameter and A_{22} is a stable matrix. Setting $\epsilon=0$ in (2.11) results in

$$\dot{x}_1(t)=(A_{11}-A_{12}A_{22}^{-1}A_{21})x_1(t) \quad (2.12)$$

$$x_2(t)=-A_{22}^{-1}A_{21}x_1(t). \quad (2.13)$$

Note that the order of the system is reduced, since the differential equation (2.11) becomes the algebraic equation (2.13). In *singular perturbation theory*, the solution to (2.12) and (2.13) is called a zeroth order approximation to (2.10) and (2.11). Singular perturbation theory deals with the properties of zeroth and higher order approximations to systems of the form (2.10) and (2.11). We term the replacement of (2.11) by the algebraic relation (2.13) a *strong coupling approximation*.

The equation (2.12) is an approximate aggregate model [25] for (2.10) and the system described by (2.10) and (2.11). It can be shown[3] [26], [27], that n_1 (the dimension of $x_1(t)$) of the eigenvalues of (2.10), (2.11) are approximated by the eigenvalues of $(A_{11}-A_{12}A_{22}^{-1}A_{21})$, and the other n_2 (the dimension of $x_2(t)$) eigenvalues are approximated by the eigenvalues of A_{22}/ϵ. The latter fast modes are discarded in the singular perturbation approach.

The basic idea expressed by (2.10) to (2.11) has been exploited by many researchers. We discuss here only the work that seems to us particularly relevant to decentralized control of large scale systems.

One reason for the interest in singular perturbation methods is of course the possible reduction in computation associated with dealing with reduced order models. For example, in [21], [28], [29], the optimal linear-quadratic regulator problem for a singularly perturbed system is solved. It is demonstrated that the zeroth order solution can be improved by the method of matched asymptotic expansions [29].

As another example, consider linear systems $S_{\epsilon\mu}$ of the form

$$S_{\epsilon\mu}\left\{\begin{bmatrix}\dot{x}(t)\\ \epsilon\dot{y}(t)\\ \mu\dot{z}(t)\end{bmatrix}=\begin{bmatrix}A_{11}&A_{1H}&A_{1L}\\ A_{H1}&A_{HH}&A_{HL}\\ A_{L1}&A_{LH}&A_{LL}\end{bmatrix}\begin{bmatrix}x(t)\\ y(t)\\ z(t)\end{bmatrix}\right. \quad (2.14)$$

The states $x(t)$, $y(t)$, and $z(t)$ correspond to the normal, high- and low-frequency parts of $S_{\epsilon\mu}$, respectively. Define the *mid-frequency* approximation

$$S_{0\infty}\left\{\begin{array}{l}\dot{x}(t)=A_{11}x(t)+A_{1H}y(t)\\ 0=A_{H1}x(t)+A_{HH}y(t)\end{array}\right. \quad (2.15)$$

[3]The terminology weak coupling is employed in [26]. However, the problem studied is the approximation of (2.10), (2.11) by (2.12), (2.13).

the *high-frequency* approximation

$$S_H:\epsilon\dot{y}(t)=A_{HH}y(t) \quad (2.16)$$

and the *low-frequency* approximation

$$S_L\left\{\begin{array}{l}0=A_{11}x(t)+A_{1H}y(t)+A_{1L}z(t)\\ 0=A_{H1}x(t)+A_{HH}y(t)+A_{HL}z(t)\\ \mu\dot{z}=A_{L1}x(t)+A_{LH}y(t)+A_{LL}z(t)\end{array}\right. \quad (2.17)$$

to $S_{\epsilon\mu}$. Each of these approximations might be of interest to an analyst depending on the application of interest.

In [30], it is shown that if all the linear systems $S_{0\infty}$, S_H, and S_L are stable, then there exists $\epsilon_0>0$ and $\mu_0>0$ such that $S_{\epsilon\mu}$ is asymptotically stable for any $\epsilon\in[0,\epsilon_0]$ and any $\mu\in[\mu_0,\infty]$. This result permits a stability test on a series on reduced order problems.

Perhaps an even more significant reason for interest in singular perturbation theory is the simplified controller structures obtainable ([31]–[35]). To illustrate this point, we will reinterpret the analysis of [31] as a result in multilevel filtering. Suppose a system is given of the form

$$\dot{x}_1(t)=A_{11}x_1(t)+A_{12}x_2(t)+A_{13}x_3(t)+\xi_1(t) \quad (2.18)$$

$$\epsilon\dot{x}_2(t)=A_{21}x_1(t)+A_{22}x_2(t)+\xi_2(t) \quad (2.19)$$

$$\epsilon\dot{x}_3(t)=A_{31}x_1(t)+A_{33}x_3(t)+\xi_3(t) \quad (2.20)$$

with observations

$$y_2(t)=C_{21}x_1(t)+C_2x_2(t)+\theta_2(t) \quad (2.21)$$

$$y_3(t)=C_{31}x_1(t)+C_3x_3(t)+\theta_3(t) \quad (2.22)$$

where $\xi_1(t)$, $\xi_2(t)$, $\xi_3(t)$, $\theta_2(t)$, $\theta_3(t)$ are independent white noise processes. Suppose, moreover, that the matrices A_{22} and A_{33} are stable, and that the eigenvalues of A_{22}/ϵ and A_{33}/ϵ have negative real parts with magnitudes much larger than the magnitudes of the eigenvalues of A_{11}.

Setting $\epsilon=0$ gives the low-order approximation

$$\dot{x}_1(t)=(A_{11}-A_{12}A_{22}^{-1}A_{21}-A_{13}A_{33}^{-1}A_{31})x_1(t)+\xi_1(t)-A_{22}^{-1}\xi_2(t)-A_{33}^{-1}\xi_3(t) \quad (2.23)$$

$$y_2(t)=(C_{21}-C_2A_{22}^{-1}A_{21})x_1(t)-A_{22}^{-1}\xi_2(t)+\theta_2(t) \quad (2.24)$$

$$y_3(t)=(C_{31}-C_3A_{33}^{-1}A_{31})x_1(t)-A_{33}^{-1}\xi_3(t)+\theta_3(t). \quad (2.25)$$

Let $\hat{\hat{x}}_1(t)$ denote the filter estimate of $x_1(t)$ corresponding to (2.23)–(2.25). Then, since $x_1(t)$ changes very slowly relative to $x_2(t)$ and $x_3(t)$, it would seem that filters for subsystems 2 and 3 could be constructed by considering x_1 to be constant in (2.19)–(2.22) with corresponding estimate $\hat{\hat{x}}_1$. Thus, the filter should be of the form

$$\hat{\hat{x}}_2(t)=-A_{22}^{-1}A_{21}\hat{\hat{x}}_1(t)+\hat{\eta}_2(t) \quad (2.26)$$

$$\epsilon\dot{\hat{\eta}}_2(t)=A_{22}\hat{\eta}_2(t)+K_2[C_2\hat{\eta}_2(t)+C_{21}\hat{\hat{x}}_1(t)-y_2(t)] \quad (2.27)$$

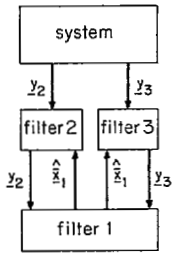

Fig. 2. Hierarchical filter configuration.

$$\hat{\bar{x}}_3(t) = -A_{33}^{-1}A_{31}\hat{\bar{x}}_1(t) + \hat{\eta}_3(t) \tag{2.28}$$

$$\epsilon\dot{\hat{\eta}}_3(t) = A_{33}\hat{\eta}_3(t) + K_3\left[C_3\hat{\eta}_3(t) + C_{31}\hat{\bar{x}}_1(t) - y_3(t)\right]. \tag{2.29}$$

This filtering configuration is illustrated in Fig. 2.

Notice the hierarchical structure. Moreover, since subsystem 1 has slower dynamics than subsystems 2 and 3, its estimate can be updated much less frequently than those of subsystems 2 and 3 (i.e., a numerical integration algorithm with a larger step size can be employed).

It can be shown [31] that this intuitive filter design has the property that

$$\hat{\bar{x}}_1 \to \hat{x}_1, \quad \hat{\bar{x}}_2 \to \hat{x}_2, \quad \hat{\bar{x}}_3 \to \hat{x}_3$$

in quadratic mean as $\epsilon \to 0$, where $\hat{x}_1$, $\hat{x}_2$, $\hat{x}_3$ denote the *optimal* estimates.

It is felt that this result is quite significant. Note that *no* assumption of a hierarchical structure was made at the outset, only that the system under control had a certain natural two-time scale structure. The mathematical analysis then indicates that the hierarchically structured filter is close to the optimal filter, where optimality is measured in terms of mean square estimation error alone. Of course, a more refined notion of optimality taking into account the cost of implementing the filter might well establish the superiority of the two time scale filter.

As demonstrated, singular perturbation methods offer a great degree of promise for model simplification, and for obtaining hierarchically structured filters and controllers. However, from a practical point of view, the main problem with these methods is that a model of a physical system is hardly ever given in the form of (2.10) and (2.11), with the slow and fast variables separated and the parameter ϵ conveniently appearing in the left-hand side of the equations. It is a completely nontrivial exercise, requiring considerable physical insight, to model a physical system with slow and fast modes in the framework demanded by theory. This problem is naturally more severe for a poorly understood large scale system. The development of *systematic* procedures for appropriately modeling large scale systems with slow and fast dynamics has not received the attention it deserves; see however [36].

An area that has not received any attention to the best of our knowledge is the development of singular-perturbation type results for discrete time systems. Many naturally discrete-time systems arise in economic models, and economic systems are indeed characterized by fast and slow dynamics so that singular perturbation techniques can be potentially useful.

Another related area, that has not received theoretical attention, is the use of sampled-data control of continuous time systems described by fast and slow modes. From an intuitive point of view, it is clear that lower sampling rates can be tolerated for estimation and control of the slow system as compared to those of the fast system. This leads naturally to *multirate* sampling, used in many physical applications. It would be very useful to use singular perturbation ideas to put such multirate sampled data designs in a theoretical perspective and in relation to hierarchical control.

III. Stability of Interconnected Systems

A. Introduction

The literature on qualitative properties of large systems is almost entirely limited to the study of stability. There is some important work on controllability which will be reviewed in Section IV. Also of relevance are isolated papers dealing with guaranteeing a differential equation representation (and uniqueness of its solution) for an interconnected system starting with such a representation for the constituent subsystems—see [37]–[38], and [39]—that discusses the order, minimality, etc., of the representation for the interconnected system. Cascade connections are discussed in [40], and in [41] feedback connections are treated also.

Here we review exclusively the literature on stability; with few exceptions these studies proceed in the following manner.

Step 1: The system is supposed to consist of interconnected subsystems. It is assumed that this decomposition or tearing has already been specified, and that a description of each subsystem and a description of the interconnection is available.

Step 2: It is assumed that each subsystem, when considered in isolation, is stable. Furthermore, some quantitative measure of this stability (e.g., a lower bound on the rate of decrease of a Lyapunov function) is available.

Step 3: A condition is now specified in terms of this quantitative measure and some quantitative measure of the magnitude of the interconnection, and it is shown that the interconnected system is stable if the condition holds.

By virtue of Step 2 we know that in the absence of any coupling of the subsystems, the collection of isolated subsystems is stable. By continuity we can expect that if the coupling is sufficiently small or weak, then the interconnected system will be stable. Two consequences follow. First, in comparing different results, a major criterion will be how large is the permitted magnitude of the coupling. Second, there are many important practical examples of interconnected systems in which stability arises

precisely because of strong coupling. It follows that systems with this property are not usefully studied by the techniques under review.

Let us turn to more technical considerations. There are two approaches to stability analysis: the Lyapunov method, e.g., [42], and the input-output method, e.g., [43].

B. Lyapunov Methods

Step 1: The overall system

$$\dot{x} = F(x,t) \tag{3.1}$$

is assumed to consist of isolated subsystems

$$\dot{z}_i = f_i(z_i,t) + g_i(u,t), \qquad i = 1,\cdots,k \tag{3.2}$$

interconnected by the constraint

$$u(t) = x(t). \tag{3.3}$$

Here $x \in R^n$, $z_i \in R^{n_i}$ and under (3.1), (3.2), and (3.3)

$$x'(t) = (z_1'(t)\cdots z_k'(t)). \tag{3.4}$$

Often it is assumed that the interaction term g_i in (2.2) is *separable*, see Fig. 3., i.e.,

$$g_i(x,t) = \sum_j g_{ij}(z_j,t). \tag{3.5}$$

Finally, assume that $\mathbf{0}$ is an equilibrium state, i.e.,

$$f_i(\mathbf{0},t) \equiv \mathbf{0}, \quad g_i(\mathbf{0},t) \equiv \mathbf{0}, \quad F(\mathbf{0},t) \equiv \mathbf{0}. \tag{3.6}$$

Step 2: Suppose there are functions V_i, d_i, and w_i such that, along trajectories of (3.2), we have

$$\dot{V}_i(z_i(t),t) \leqslant -d_i(z_i(t),t) + w_i(z_i(t),u(t),t), \tag{3.7}$$

where V_i is PDU (positive-definite, decrescent and radially unbounded).[4] If V_i is continuously differentiable then taking

$$d_i(z_i,t) = -\left(\frac{\partial V_i}{\partial t}(z_i,t) + \frac{\partial V_i}{\partial z_i}(z_i,t) f_i(z_i,t)\right);$$

$$w_i(z_i,u,t) = \frac{\partial V_i}{\partial z_i}(z_i,t)\, g_i(u,t), \tag{3.8}$$

gives equality in (3.7).

Now suppose that the right-hand side of (3.7) can be bounded

$$\begin{aligned}\dot{V}_i(z_i(t),t) &\leqslant -d_i(z_i(t),t) + w_i(z_i(t),u(t),t)\\ &\leqslant h_i(\gamma_1(z_1(t),t),\cdots,\gamma_k(z_k(t),t))\end{aligned} \tag{3.9}$$

where the γ_i are also PDU. Note that for the procedure to be useful it will be necessary to obtain γ_i in terms of V_i and h_i in terms of the magnitude of the interaction signal w_i.

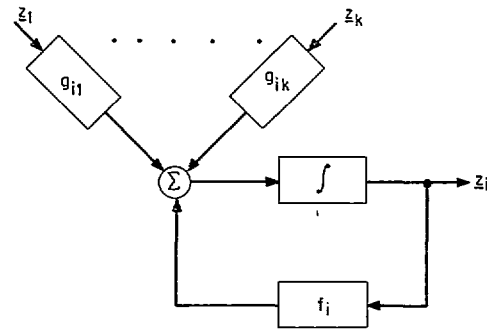

Fig. 3. Separable interaction terms.

Step 3: It is now necessary to place conditions on the h_i, γ_j so that $V_i(z_i(t),t) \to 0$ as $t \to \infty$. Two ways of doing this may be distinguished; the vector Lyapunov method and the weighted-sum method.

The vector Lyapunov method is well discussed in [44]. For example, one of the results in [44] is that if (3.9) holds with $\gamma_i(z_i,t) = V_i(z_i,t)$ and with $h_i(\gamma_1,\cdots,\gamma_k,t)$ nonincreasing in γ_j for $j \neq i$, and if the origin is an asymptotically stable equilibrium of the k-dimensional system

$$\dot{v}_i = h_i(v_1,\cdots,v_k,t), \qquad i = 1,\cdots k, \tag{3.10}$$

then the overall system (3.1) is asymptotically stable. The k-vector $V'(x,t) = (V_1(z_1,t),\cdots,V_k(z_k,t))$ is called a vector Lyapunov function. Now in all the examples that we know of in which this result is used including the original work of [45] and later work, [46]–[53], (3.9) takes the form of the inequality

$$\dot{V}_i(z_i(t),t) \leqslant -\sum_j a_{ij}\phi_j(\|z_j(t)\|) \leqslant -\sum_j \tilde{a}_{ij} V_j(z_j,t) \tag{3.11}$$

where the ϕ_i are PDU and where A, $\tilde{A}$ are matrices with positive diagonal and nonpositive off-diagonal entries. Hence, the a_{ij} or $\tilde{a}_{ij}$, $j \neq i$, give a bound on the magnitude of the subsystem interactions. Also, observe that in isolation ($a_{ij} = \tilde{a}_{ij} = 0$, $i \neq j$) the subsystems are stable since $a_{ii} > 0$, $\tilde{a}_{ii} > 0$. From (3.11) we see that (3.10) takes on the special form of the linear differential equation

$$\dot{V} = -\tilde{A} V. \tag{3.12}$$

It is well-known that (3.12) is stable if and only if $\tilde{A}$ is an M-matrix.[5] Step 3 of the procedure can now be completed by specifying conditions on the $\tilde{A}$ matrix which guarantee that it is an M-matrix. For example, if $\tilde{A}$ is diagonally dominant, i.e., $\tilde{a}_{ii} > \sum_{j \neq i} |\tilde{a}_{ij}|$, then it is an M-matrix.

[4] V_i is PDU if and only if $\phi_1(\|z_i\|) \leqslant V_i(z_i,t) \leqslant \phi_2(\|z_i\|)$ where the ϕ_i satisfy $\phi_i(0) = 0$, ϕ_i continuous and strictly increasing, and $\phi(\infty) = \infty$.

[5] A matrix T with positive diagonal and nonpositive off-diagonal entries is an M or Metzler matrix if the following equivalent conditions hold: i) T^{-1} exists and all its entries are nonnegative, ii) the leading principal minor determinants of T are all positive, iii) the eigenvalues of T have positive real parts. For a concise discussion see [54].

In the cited papers which use the vector Lyapunov method the matrix $\tilde{A}$ is an M-matrix only if A is. But then, as shown below, stability of (3.1) can also be proved using the weighted-sum approach. It therefore appears that as far as applications are concerned, the weighted-sum Lyapunov function approach is at least as good as the vector Lyapunov approach.

In the weighted-sum approach one proceeds using the following obvious results. Suppose (3.9) is such that there exist positive numbers $\alpha_1, \cdots, \alpha_k$ so that $-\Sigma\alpha_i h_i(\gamma_1, \cdots, \gamma_k, t)$ is PDU with respect to $\gamma = (\gamma_1, \cdots, \gamma_k)$. Let $V(x,t) = \Sigma\alpha_i V_i(z_i, t)$. Then along the trajectories of (3.1), $\dot{V}(x(t),t) < 0$ for $x(t) \neq 0$ and (3.1) is asymptotically stable.

Particular applications of this have been made for the special case where, for PDU ϕ_i,

$$\gamma_i(z_i, t) = \phi_i(\|z_i\|). \tag{3.13}$$

The usual way in which the ϕ_i of (3.13) are obtained is to bound the isolated system behavior (3.9) as

$$\phi_i(\|z_i\|) \leqslant d_i(z_i(t), t) \tag{3.14}$$

so that we have

$$\dot{V}_i(z_i(t), t) \leqslant -\phi_i(\|z_i\|) + w_i(z_i, u, t). \tag{3.15}$$

This is the procedure used in [54]–[59]. For example, we have the following result [57], [59].[6] Suppose in (3.10) we have $h_i(\phi_1, \cdots, \phi_k, t) = -\Sigma a_{ij}\phi_j$ where A is a constant matrix. If A is an M-matrix, then there exist positive $\alpha_1, \cdots, \alpha_k$ such that $-\Sigma a_i h_i$ is PDU, hence, (3.1) is asymptotically stable.

Another application of the weighted sum approach suitable for quadratic Lyapunov functions is to require that there exist positive constants $\alpha_1, \cdots, \alpha_k$ so that

$$-\Sigma\alpha_i h_i(\phi_1 \cdots \phi_k, t) = -\phi' B \phi \tag{3.16}$$

where $\phi = (\phi_1, \cdots, \phi_k)$ and B is positive definite [59].

While the literature concerned with the use of weighted-sum Lyapunov techniques is large, a fairly comprehensive coverage is given in [59], where many variations are developed for systems described by ordinary differential equations, difference equations, sampled-data systems, and functional differential equations. We may note that most results are such that the guarantee connective stability [51] in the sense that stability is maintained even if one or more of the g_{ij} in (3.5) vanish.[7]

The papers described so far have arrived at a simplification of the test for Lyapunov stability. A potentially important advantage of these tests is their decomposition into two stages.

1) Behavior of each isolated subsystem is characterized by its own Lyapunov function and the characterization does *not* require the knowledge of other subsystem models.

2) The stability test of the interconnected system (carried out at a 'higher' level) is in terms of the individual Lyapunov functions and bounds on interconnections. Thus, detailed knowledge of the system structure is not needed, but only (in a sense) aggregate knowledge [61].

However there are some severe drawbacks.

1) Tests are successful only when the coupling among subsystems is weak. This occurs in two ways: first, the isolated subsystems are assumed to be stable and at the equilibrium of the overall system all interactions vanish (see (3.6)) and second, almost all tests perform better when the interactions are weaker.[8]

2) The overall system is analyzed by partitioning it into disjoint subsystems. In many interesting problems it is more natural to consider "overlapping" subsystems. (See [62] for an indication of how this idea might be carried out.)

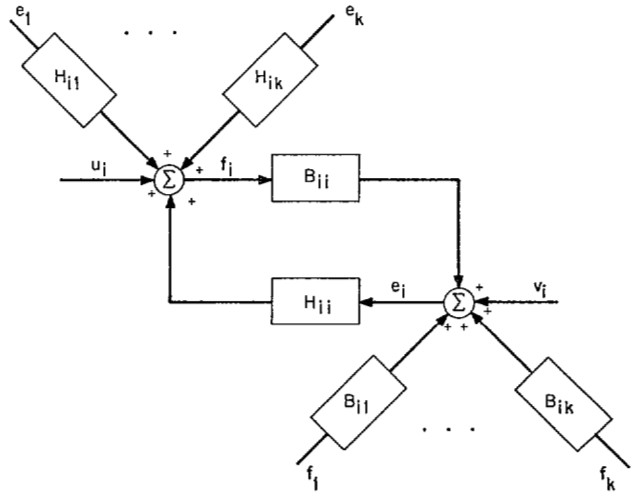

Fig. 4. Input–output relations for subsystem i.

C. Input–Output Methods

Step 1: The overall system (Fig. 4) is assumed to consist of interconnected subsystems whose inputs and outputs are related by

$$\left.\begin{aligned} u_i &= f_i - \sum_{j=1}^{k} H_{ij} e_j \\ v_i &= -\sum_{j=1}^{k} B_{ij} f_j + e_i \end{aligned}\right\} \quad i = 1, \cdots, k \tag{3.17}$$

[6]It should be noted that the authors of [57] describe their approach as a vector Lyapunov function; actually the weighted-sum Lyapunov function approach is employed.

[7]To a considerable extent this advantage is lost by the implicit assumption that the equilibrium state remains unchanged when the isolated systems are interconnected (see (3.6)).

[8]This is practically built into the procedures. Thus, for example, in (3.11) we see that the effect of interactions is bounded only in terms of their magnitudes. It makes no difference whether these interactions are stabilizing or destabilizing. The tests then will necessarily be very conservative.

where u_i, e_i, f_i may be all vector valued functions. The input of (3.17) is $(u_1,\cdots,u_k,v_1,\cdots,v_k)$ and its output is $(e_1,\cdots,e_k,f_1,\cdots,f_k)$; H_{ij} and B_{ij} are nonlinear, nonanticipative operators on $L_{pe}[0,\infty)$.[9] Note that the output is given implicitly as a function of the input. An appropriate notion of stability in this description is the one of boundedness.[10]

Step 2: Suppose that each of the operators H_{ij}, B_{ij} has finite gains $g(H_{ij})$, $g(B_{ij})$[11] and let $\boldsymbol{G}(H)$, $\boldsymbol{G}(B)$ be the $k\times k$ matrices with these gains as entries. Note that the off-diagonal entries of $\boldsymbol{G}(H)$, $\boldsymbol{G}(B)$ give a bound on the magnitude of the interactions.[12] [For purposes of application it is important to note that if the isolated subsystems H_{ij}, B_{ij} are either memoryless or time invariant then these gains can be readily computed from the graphs of the nonlinearities or from the Nyquist plots of linear systems (see [63] and [43, ch. 5]).

Step 3: It is now necessary to place conditions on $\boldsymbol{G}(H)$, $\boldsymbol{G}(B)$ so that the overall system (3.17) has bounded gain. One such result, proved in [64], is that if $\boldsymbol{I}-\boldsymbol{G}(B)\boldsymbol{G}(H)$ is an M-matrix, then (3.17) has finite gain.

Analogous conditions may be obtained for guaranteeing finite incremental gains and for continuity. Simpler versions of this result have been given in [65] and [66]. As in the Lyapunov method the conservative test proposed by this result guarantees connective stability, i.e., if $\boldsymbol{I}-\boldsymbol{G}(B)\boldsymbol{G}(H)$ is an M-matrix, so is $\boldsymbol{I}-\boldsymbol{G}(B')\boldsymbol{G}(H')$, where $g(B'_{ij})\leqslant g(B_{ij})$ and $g(H'_{ij})\leqslant g(H_{ij})$.

For the special case where $H_{ij}=0$ for $i\neq j$ some additional refinements are possible. With more restrictive conditions on the isolated subsystems, Lasley and Michel [64] obtain Popov-type stability conditions, and Araki [67] obtains a circle stability criterion. The latter extends previous work in [65]–[66] and [69]–[70].

The preceding result is reminiscent of the "loop gain less than one" criterion of Zames [63]. For the case where L_{pe} is L_{2e} it should also be possible to use the properties of positive operators in the way Zames does. Thus, for example, consider $H=\{H_{ij}\}$, $B=\{B_{ij}\}$ as operators on L_{2e}. Then we may conclude from Zames' positivity results that (3.17) has finite L_2 gain provided that H is positive, $-B$ is strongly positive and the $g(B_{ij})$ are finite. To get an analogous result it would be necessary to guarantee positivity of both H, and $-B$ in terms of conditions on the subsystems. The case where H is diagonal and B is a matrix of constant gains is treated in [71].

In evaluating the work on Lyapunov stability of interconnected systems we have already noted the conservative nature of the stability tests; for the same reasons the input–output stability tests for interconnected systems also tend to be conservative. In comparing the two classes of tests it seems to us that the input-output stability tests are superior for three reasons. First of all, the input–output methods are often less conservative and substantially easier to apply than Lyapunov methods; this has been convincingly demonstrated in [65]. Second, since the various gains are more readily related to design parameters than are Lyapunov functions, the input–output tests lead to relatively simple design guidelines (see [64] and [69]). Third, as we observed earlier the Lyapunov approach is limited by the assumption that the equilibrium is known or that the isolated subsystem equilibria coincide with that of the interconnected system equilibrium. The incremental gain version of the input–output results is, however, unaffected by changes in the equilibria.

D. Other Results

The two main classes of stability results continue to be unrelated in spite of some key suggestive remarks of Willems [72]. Some very recent work by Safonov [73] clarifies the relationship between the Lyapunov and input–output methods; however, its implications with regard to interconnected systems are only partially developed.

Callier, Chan, and Desoer [74] give an algorithm, based on the graph of the system interconnections, for relabeling the subsystems in such a way that the computation of input–output stability conditions is simplified. Lyapunov stability conditions for interconnected systems subject to random disturbances are obtained in [75] and [76]. Local Lyapunov stability conditions are given in [47] and [59]. Trajectory bounds are given in [77] and [78]. We have already noted that a singularly perturbed system can be regarded as consisting of a fast system tightly coupled to a slow one, and that the stability of the overall system can be guaranteed by separate conditions on the degenerate and boundary layer systems. Grujic [79] considers several interconnected singularly perturbed systems.

Finally, we note the very recent appearance of a monograph by Michel and Miller [80] that deals with several topics discussed in this section.

IV. Decentralized Feedback Control

A. Introduction

Research in decentralized control has been motivated by the inadequacy of conventional modern control theory to deal with certain issues of concern in large scale systems. A key concept in modern control theory is that of state feedback. By techniques such as linear-quadratic (LQ) optimal control or pole placement it is possible to achieve improved system behavior by using state feedback. However, it is often impossible to instrument a system to the extent required for full state feedback, so

[9] For any function $x(t)$ defined on $[0,\infty)$ and $T<\infty$ let $P_T x$ be the function $(P_T x)(t)=x(t)$, $t\leqslant T$ and $(P_T x)(t)=0$, $t>T$. $L_{pe}[0,\infty)$ consists of all functions x such that $\|P_T x\|_p=\int_0^\infty\|(P_T x)(t)\|^p dt<\infty$ for all $T<\infty$. An operator F on $L_{pe}[0,\infty)$ is nonanticipative if $P_T F P_T=P_T F$ for all T.

[10] An operator F on $L_{pe}[0,\infty)$ is bounded if for each $r_1<\infty$ there is $r_2<\infty$ such that $\|Fx\|_p<r_2$ whenever $\|x\|_p<r_1$.

[11] The gain of F is $g(F)=\sup\{\|P_T Fx\|_p/\|P_T x\|_p \mid T<\infty,\ x\in L_{pe}$ and $\|P_T x\|_p\neq 0\}$.

[12] Again note the conservative bias which is introduced by considering only the magnitude.

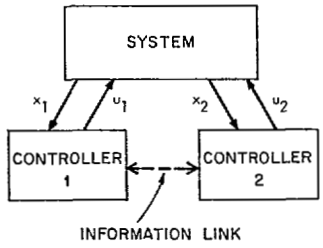

Fig. 5. Partial decentralization.

techniques ranging from linear-quadratic-Gaussian (LQG) control to observer-based control to time domain compensator design techniques have evolved to overcome this difficulty. However, a key characteristic of all of these techniques is that a design results in which every sensor output affects every actuator input. We term this situation centralized control. Of course, in many systems, but particularly large scale systems, it is impossible to incorporate so many feedback loops into the design.

Decentralized control theory has risen in response to this difficulty. The basic characteristic of decentralized control is that there are restrictions on information transfer between certain groups of sensors or actuators. For example, consider Fig. 1, where only the state variables $\boldsymbol{x}_1$ are used to form the control $\boldsymbol{u}_1$, and similarly only the state variables $\boldsymbol{x}_2$ are used to form the control $\boldsymbol{u}_2$. This is a case of total decentralization. However, intermediate restrictions on information transfer between the controllers (Fig. 5) are also possible. These cases in which the rate of information transfer is constrained to be below that necessary for implementation of the centralized solution, are termed partial decentralization.

Decentralized control is thus intimately concerned with feedback. Therefore, linear problems are those most often studied, since nonlinear feedback control theory is much less developed than the linear theory. If the feedback is to be obtained by solution of an optimization problem, a stochastic problem formulation is often assumed since the optimal open-and closed-loop solutions are identical for deterministic problems. It is important to note that the term decentralization refers to the control structure implemented; the control laws may be designed in a completely centralized way. The reader should contrast the decentralized control problem with the hierarchical control problem to be discussed in Section V.

B. Decentralized Stochastic Control Methods

In centralized control, one viable approach to design of feedback control laws for time invariant linear systems is by minimization of an infinite horizon quadratic performance index. In the deterministic full state feedback (LQ) case, this can be accomplished by solution of an algebraic Riccati equation, for which a variety of algorithms are available. As is well known, a stable design is always produced (under reasonable and rather weak conditions) for any choice of the performance index, and by varying the performance index the designer can trade off between the error in driving the state to zero after an impulsive disturbance and the corresponding control activity (both measured in an integral squared error sense). Less often emphasized, but perhaps even more important from an applications point of view, the LQ design method allows asymptotic pole placement by appropriate choice of performance index [81], has excellent sensitivity and robustness properties according to various criteria including the classical gain and phase margins [82], and handles multiloop problems in exactly the same framework as single-loop problems. These results generally extend to the stochastic (LQG) case, in which disturbances are modeled by passing white Gaussian noise through finite dimensional shaping filters, and in which only partial state information is available. However, the separation theorem solution of the LQG problem requires that a Kalman filter be implemented to reconstruct the missing state variables. Thus, an additional filter Riccati equation must be solved, the modes of the filter will appear in the closed-loop response of the system, and gain and phase margins will be reduced so that an LQG design will be more sensitive and less robust than a corresponding LQ design [73].

Given the desirable features of the LQG solution in the centralized case, it is not surprising that researchers have attempted to extend the approach to the decentralized case. What is surprising is how little progress has been made. We will attempt to explain this state of affairs in the sequel.

Consider a linear stochastic system

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\sum_{i=1}^{N}\boldsymbol{B}_i\boldsymbol{u}_i(t)+\boldsymbol{\xi}(t) \quad \text{(state equation)} \tag{4.1}$$

$$\boldsymbol{y}(t)=\boldsymbol{C}\boldsymbol{x}(t)+\sum_{i=1}^{N}\boldsymbol{D}_i\boldsymbol{u}_i(t)+\boldsymbol{E}\boldsymbol{\theta}(t) \quad \text{(measurement equation)} \tag{4.2}$$

in which $\boldsymbol{\xi}(t)$, $\boldsymbol{\theta}(t)$ are independent white Gaussian processes. Note that in (4.1), the total set of inputs to the system is decomposed into inputs $\boldsymbol{u}_i(t)$, each of which is assumed to be specified by a different controller. The quantity $\boldsymbol{y}(t)$ in (4.2) corresponds to all the sensor outputs of the system. The portion of these measurements available to a given controller is specified by the *information pattern* of the problem [83]. For the problem considered here the information pattern is a set of matrices

$$\mathcal{H}=\{\boldsymbol{H}_1,\boldsymbol{H}_2,\cdots,\boldsymbol{H}_N\} \tag{4.3}$$

with the interpretation that

$$\boldsymbol{z}_i(t)=\boldsymbol{H}_i(t)\boldsymbol{y}(t), \qquad i=1,\cdots,N \tag{4.4}$$

is the portion of the measurements available to controller i. The usually studied case, termed the *classical information pattern*, has

$$\boldsymbol{H}_i=\boldsymbol{I}, \qquad i=1,\cdots,N \tag{4.5}$$

so that all controllers have identical information.

Suppose that a quadratic performance index of the form

$$J=\lim_{T\to\infty}\left\{\frac{1}{T}\int_0^T\left[x'(t)\boldsymbol{Q}x(t)+\sum_{i=1}^{N}u_i'(t)\boldsymbol{R}_iu_i(t)\right]dt\right\} \tag{4.6}$$

is given. Because of the nonclassical information pattern the controls are restricted to be of the form

$$u_i(t)=\gamma_i(z_i^t) \tag{4.7}$$

where

$$z_i^t=\{z_i(\tau):0\leqslant\tau\leqslant t\}. \tag{4.8}$$

That is, the control of the ith controller is to be chosen as an arbitrary functional of the past and present information available to it, subject only to mild technical restrictions on the functions $\gamma_i(\cdot)$ to assure that the problem is well posed (see, for example, [84]).

Note that once a set of control laws $\gamma_i(\cdot)$, $i=1,\cdots,N$, is given, (4.7) can be used to eliminate the $u_i(t)$ from (4.1), (4.2), and (4.6) so that the state and observations as stochastic processes, and the cost as a random variable, become well-defined. Clearly, all these quantities depend on the choice of the γ_i. In particular the expected cost, denoted

$$\mathcal{J}(\gamma_1,\cdots,\gamma_N)=EJ \tag{4.9}$$

depends on the choice of control law. The stochastic optimal control problem then is the functional minimization of $\mathcal{J}(\gamma_1,\cdots,\gamma_N)$.

In the case of the classical information pattern, the well known separation theorem asserts the optimality of a linear feedback of state estimates generated by Kalman filter. There are two aspects of this solution to be emphasized. First, the optimal control law is *linear* and, therefore, is at least as good as any *nonlinear* control law. Second, compared to a control law which can depend on the *infinite dimensional* record of past measurements, the optimal control law depending only on the *finite dimensional* Kalman filter state estimates is superior. Both of these properties fail in the decentralized case characterized by a nonclassical information pattern.

The possibility of nonlinear solutions to nonclassical LQG stochastic control problems first became apparent in 1968. Witsenhausen [85] formulated an apparently trivial, two-stage, scalar, discrete time LQG problem and showed that in a limiting case an ad hoc nonlinear control law outperformed the best linear control law. The optimal solution to this Witsenhausen counterexample has never been found, although considerable insight has been obtained into the nature of the solution.

To understand the reason for nonlinear solutions in decentralized LQG problems, we will consider a rather degenerate case, termed the *control-sharing* information pattern [86] to [88]. In this problem, the matrices $\boldsymbol{C}$, $\boldsymbol{D}_1,\cdots,\boldsymbol{D}_N$, $\boldsymbol{H}_1,\cdots,\boldsymbol{H}_N$, $\boldsymbol{E}$ are specified in such a way that each $z_i(t)$ contains $u_j(t)$ (assumed measured and communicated without noise) for $j\neq i$. In other words, each controller knows exactly and instantaneously the controls generated by all other controllers. To see that a nonlinear solution is indeed optimal, suppose that at time t controller i would apply the optimal control

$$u_i^*(t)=\tfrac{1}{3}=.3333\cdots \tag{4.10}$$

if the information pattern were classical. Suppose that his observation at time t is

$$z_i(t)=\tfrac{2}{3}=.666\cdots. \tag{4.11}$$

Consider the consequences of the choice of control

$$u_i(t)=.333\cdots336666\cdots. \tag{4.12}$$

Clearly, the control specified in (4.12) is an attempt by controller i to *signal* the value of his observation to the other controllers which in turn can "decode" $z_i(t)$ from $u_i(t)$. Moreover, the control $u_i(t)$ can be made arbitrarily close to $u_i^*(t)$ so that the substitution of $u_i(t)$ for $u_i^*(t)$ affects the system response by an infinitesimaly small amount. If all the controllers cooperate and each employs such a scheme, an expected cost arbitrarily close to the cost for the corresponding problem with classical information pattern can be achieved. Of course, these signaling control laws are quite nonlinear.

The argument of the preceding paragraph is reminiscent of the argument that a noiseless information channel has an infinite capacity. This is not coincidental; many communication problems, in which the information available to the device controlling a channel input must differ from that of the device observing the channel output for nontriviality, can be regarded as nonclassical stochastic control problems. However, in nonclassical stochastic control choice of an input has a *dual* purpose; communication through the system dynamics and sensors to the other controllers and direct control of the system. Thus, interleaving the digits of $u_i^*(t)$ and $z_i^*(t)$ permits both of these quantities to be communicated perfectly by transmission of a single real variable through a noiselsss channel. In contrast, for the control-sharing problem, the digits of $z_i(t)$ must be added at the end of a truncated version of $u_i^*(t)$ so that $u_i(t)$ and $u_i^*(t)$ are approximately equal and have similar effects on the system. This can only be accomplished with a finite, albeit arbitrarily small, error.

From the control-sharing example, one can speculate that nonlinear solutions in nonclassical LQG control are due to the actions of the decentralized controllers as they try to signal information to one another using the control system as a communication channel. However, this notion is difficult to make precise, since it is impossible to ascertain precisely what portion of a given system input is applied to purposes of signaling and what portion for control.[13] A rigorous definition of signaling control laws

[13]A similar problem arises in adaptive control, where it is difficult to separate out exactly what portion of a control signal is being used to probe the system and what part to control it.

has been given [89], but the definition really deals with *potential* signaling.

Despite the difficulty of making precise statements, all of the work we are aware of in decentralized stochastic control supports our interpretation of the nonclassical problem. For example, it has been shown that discrete-time LQG problems with certain nonclassical information patterns have linear solutions [90]–[92]. These information patterns, termed *partially nested*, have the property that the information of any controller at a given time is unaffected by any of the control actions previously applied to the system. For such problems, there is of course no possibility of signaling. As another example, the nonlinear strategy of the Witsenhausen counterexample amounts to an effort to signal the value of the present state from one controller to another controller acting at a later time. Interestingly, this strategy is proved superior to the best linear control law only in a limiting case which corresponds to a high signal-to-noise ratio for communication through the state and observation equations. This is not surprising if one is familiar with basic communication theory in which it is well known that nonlinear modulation is optimal only in the high signal-to-noise ratio region [93]. In fact, recently a counterexample has been constructed in which a nonclassical stochastic control version of a simple communication problem is shown to possess a nonlinear solution in the high signal-to-noise ratio region [94].

One can question the desirability of using nonlinear signaling strategies to communicate through the control system dynamics on several grounds. Implementation would be exceedingly complex, and performance would seem to be quite sensitive to variations in system parameters. In any case determination of these signaling strategies has been shown to be equivalent to an infinite-dimensional, nonconvex optimal control problem with neither analytical nor computational solution likely to be forthcoming in the forseeable future [95]. This fact of life forces one to reevaluate the problem formulation. Recall that the property of the centralized solution that we are trying to generalize to the decentralized case is that the solution optimizes a Bayesian criterion of performance over the class of all functions of past observations. Happily, as we described previously, the optimal solution in the centralized case is easily computed and implemented, and has other desirable features not explicitly required in the definition of optimality. While mean square performances indices are not unimportant, the separation theorem solution would be of only academic interest if it did not possess these other desirable properties (e.g., robustness).

The preceding discussion suggests that we are perhaps trying to generalize the wrong properties of the separation theorem solution to the decentralized case. Perhaps if our insistence that the solution be optimal with respect to the very wide class of all functions of past observations is relaxed, a solution that is computable, implementable, and robust can be obtained. Of course, some increase in the quadratic cost functional must be tolerated if nonlinear signaling is not permitted. However, for cases in which the control system dynamics make a poor communication channel, (and this is usually the case for large scale systems) this loss should be minimal.

One approach that immediately suggests itself from the above discussion is to *a priori* restrict attention to linear control laws only. However, this approach runs into difficulty due to the so-called *second guessing* phenomenon [96], [97]. Specifically, in the usual LQG centralized case, the separation theorem specifies the optimal control to be a linear feedback of Kalman filter state estimates, so that the order of the optimal compensator is equal to that of the system. In the decentralized case, this result is not obtained, as illustrated by the following argument. Consider a two controller system, and suppose that one of the controllers consists of a linear feedback of state estimates. Then the other controller must, according to the usual separation theorem result, consist of a linear feedback of estimates of the system states and of estimates of the other controller's estimates. Repeated application of this argument shows that neither of the controllers can be finite-dimensional. An alternate analysis is to formulate the optimization problem in the frequency domain, where it can be shown that the optimal controller transfer functions are irrational [98].

From this discussion we can see that restricting attention to control laws that are arbitrary linear functions of past observations does not necessarily result in an optimal solution that is readily computed or implemented. A logical next attempt is to restrict the controllers to be not only linear but also of a fixed structure and optimize over the parameters of that structure. This approach was suggested first in the context of differential games [96], and has been applied to decentralized stochastic control and estimation by several authors in both finite [99] and infinite horizon settings [100], [101]. The fixed structure approach is feasible and may well be the appropriate extension of the LQG design methodology to decentralized control. However, this formulation does not address the question of what structure should be assumed. At present, it is possible to make the structural choice only on a trial-and-error basis using heuristic reasoning [102]. Obviously, engineering judgment can only help!

Most research in stochastic decentralized control beyond what we have already mentioned is concerned with determining the optimal parameters of given decentralized control structure [99]–[109]. For example, decentralized Kalman filtering is considered in [103]. Computation of the optimal gains by an ϵ-expansion technique analogous to that described in Section II is considered in [104]. A set of necessary conditions for a very general control structure is derived in [105] and [106], and solution by iterative decomposition techniques (see next section) is described. An interesting analysis of decentralized control of an infinite string of identical, linearly interconnected subsystems is found in [107]. Some application papers in this general area have appeared [108], [109].

An exception to our general statement that little attention has been paid to structural issues is the *periodic*

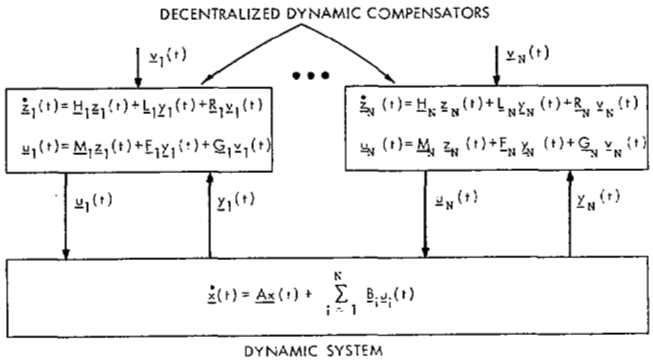

Fig. 6. Local dynamic feedback.

coordination scheme [110] (see also [111]). Although clearly a rather special situation, this work does indicate how mathematical considerations can give insight into the choice of structure.

C. Decentralized Stabilization and Pole Placement

A fundamental result in modern control theory is that the poles of a controllable linear system can be arbitrarily assigned (subject to complex pole-pairing constraints) by state variable feedback. This result has been extended to show that the poles of a closed-loop system consisting of a controllable and observable linear system with a dynamic compensator of a certain order can be freely assigned [112]. These results are of great theoretical significance, and also have served as the basis of practical synthesis procedures.

A natural generalization of the pole-placement question arises when the restriction to decentralized feedback control is made. Although several authors had looked at this question earlier, [113]–[116], the most definitive results are those of Wang and Davison [117] and Corfmat and Morse [118], [119]. In this part of the paper, we briefly summarize their results.

For a linear system the problem of decentralized pole placement can be formulated as follows. Consider the linear system

$$\dot{\boldsymbol{x}}(t)=\boldsymbol{A}\boldsymbol{x}(t)+\sum_{i=1}^{N}\boldsymbol{B}_i\boldsymbol{u}_i(t) \tag{4.13}$$

$$\boldsymbol{y}_i(t)=\boldsymbol{C}_i\boldsymbol{x}(t) \tag{4.14}$$

where $i=1,\cdots,N$ indexes the input and output variables of the various controllers. The ith controller employs dynamic compensation of the form

$$\boldsymbol{u}_i(t)=\boldsymbol{M}_i\boldsymbol{z}_i(t)+\boldsymbol{F}_i\boldsymbol{y}_i(t)+\boldsymbol{G}_i\boldsymbol{v}_i(t) \tag{4.15}$$

$$\dot{\boldsymbol{z}}_i(t)=\boldsymbol{H}_i\boldsymbol{z}_i(t)+\boldsymbol{L}_i\boldsymbol{y}_i(t)+\boldsymbol{R}_i\boldsymbol{v}_i(t) \tag{4.16}$$

as illustrated in Fig. 6. The decentralized pole-placement problem is to find matrices $\boldsymbol{M}_i$, $\boldsymbol{F}_i$, $\boldsymbol{G}_i$, $\boldsymbol{H}_i$, $\boldsymbol{L}_i$, $\boldsymbol{R}_i$ such that the closed-loop system described by (4.13)–(4.16) has prespecified poles.[14] Of course, if $[\boldsymbol{C}_i,\boldsymbol{A},\boldsymbol{B}_i]$ is controllable and observable for some i, the problem is trivial. The interesting case is to assume that (4.13) is controllable from all controls $\boldsymbol{u}_1,\cdots,\boldsymbol{u}_N$, but not from any single control $\boldsymbol{u}_i$, with a similar observability assumption.

Consider first the special case $\boldsymbol{M}_i=\boldsymbol{0}$ in the above problem. This corresponds to nondynamic decentralized output feedback. If F denotes the collection of feedback matrices $(\boldsymbol{F}_1,\boldsymbol{F}_2,\cdots,\boldsymbol{F}_N)$, then the pole-placement problem is to determine F such that the matrix

$$\boldsymbol{A}_F\triangleq\boldsymbol{A}+\sum_{i=1}^{N}\boldsymbol{B}_i\boldsymbol{F}_i\boldsymbol{C}_i \tag{4.17}$$

has a specified arbitrary set of eigenvalues. Clearly a necessary condition for pole placement in this case is that the polynomials $|\lambda\boldsymbol{I}-\boldsymbol{A}_F|$ have no common factor, i.e., that

$$\alpha(\lambda)\triangleq\text{g.c.d.}\,|\lambda\boldsymbol{I}-\boldsymbol{A}_F|=1. \tag{4.18}$$

(g.c.d. = greatest common divisor). What is much more interesting is that this condition is both necessary and sufficient for pole placement with *dynamic* decentralized compensation [117]. More generally, since the zeros of $\alpha(\lambda)$ (termed the *fixed modes* of the system) are invariant under decentralized dynamic compensation, it follows that a necessary and sufficient condition for stabilizability is that the roots of $\alpha(\lambda)$ have strictly negative real parts.

Computation of the fixed modes of a system can be accomplished as follows; see [121]. First, note that since $\boldsymbol{F}_i=\boldsymbol{0}$, $i=1,\cdots,k$, is admissible, the fixed modes of $\boldsymbol{A}$ are naturally a subset of its eigenvalues. Therefore, the first step is to compute the eigenvalues of $\boldsymbol{A}$. Second, it can be shown that if F is randomly chosen, then with probability one the fixed modes of $\boldsymbol{A}$ are common eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{A}_F$. Since algorithms are available for determining the eigenvalues of fairly high-dimensional systems (e.g., 100 states), computation of the fixed modes is reasonably straightforward.

Implicit in the proof of the pole placement result quoted above is a constructive algorithm. This algorithm requires as a first step the (possibly random) selection of F such that all the poles of $\boldsymbol{A}_F$ are distinct from those of $\boldsymbol{A}$. Then, dynamic feedback is successively employed at the control stations to place the poles that are controllable and observable from a given station. Fig. 7 illustrate this procedure [121].

Corfmat and Morse [119] have studied the decentralized feedback control problem from the point of view of determining a more complete characterization of conditions for stabilizability and pole placement. Their basic approach is to determine conditions under which a system of the form (4.13)–(4.14) can be made controllable and observable from the input and output variables of a given controller by static feedback applied by the other controllers. Then dynamic compensation can be employed at this controller in a standard way to place the poles of the system.

It is not hard to see that a necessary condition to make

[14] An extension of this problem, the decentralized servomechanism problem, has been considered by Davison [120], [121].

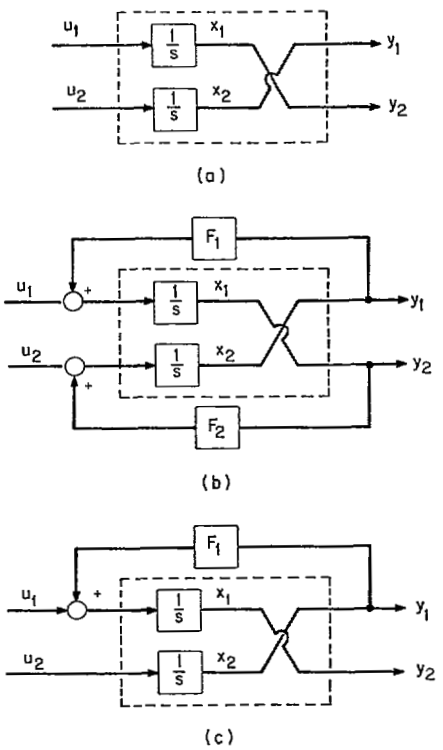

Fig. 7. Note that in the third step both modes are controllable and observable from u_2 and y_2 for all $F_1 \neq 0$.

(4.13)–(4.14) controllable and observable from a single controller is that none of the transfer functions

$$G_{ij}(s) = C_i(sI - A)^{-1}B_j \tag{4.19}$$

vanish identically. A system satisfying this condition is termed *strongly connected* [119].

If a system is not strongly connected, it is impossible to make the system controllable and observable from a single controller. In this case, it is necessary to decompose the system into a set of strongly connected subsystems, and to make each subsystem controllable and observable from one of its controllers. Thus, we can, without loss of generality, restrict our attention to the strongly connected case.

For a strongly connected system, Corfmat and Morse have given a highly interesting and rather intuitive condition that is necessary and sufficient to make (4.13)–(4.14) controllable and observable from a single controller.

They have shown that if a strongly connected system can be made controllable and observable from a single controller, then it can be made controllable and observable from *any* controller, and a necessary and sufficient for this to occur is that the system be *complete*. Completeness of a system of the form (4.13)–(4.14) is defined in terms of the transmission polynomials of the system [122]. For example, in case $N=2$, a test for completeness is [118], [119], that for all s

$$\operatorname{rank}\begin{bmatrix} (sI-A) & B_1 \\ C_2 & 0 \end{bmatrix} \geq n \tag{4.20}$$

$$\operatorname{rank}\begin{bmatrix} (sI-A) & B_2 \\ C_1 & 0 \end{bmatrix} \geq n. \tag{4.21}$$

Since

$$\operatorname{rank}(sI-A) \geq n \tag{4.22}$$

except when s is an eigenvalue of A, it is only necessary to check the above condition on the spectrum of A. Moreover, the values of s for which the matrices in (4.20) and (4.21) have less than full rank are the transmission zeros (see [123] and the references cited therein) of (C_2, A, B_1) and (C_1, A, B_2). Thus, lack of completeness is associated with pole-zero cancellations. Corfmat and Morse have also characterized, in the case of an incomplete system, the fixed modes of Wang and Davison in terms of transmission zeros.

As a practical design method, the Corfmat-Morse approach suffers some defects. First, note that even if all the modes of a large scale system can be made controllable and observable from a single controller (or a few controllers if the system is not strongly connected), some of the modes may be very weakly controllable or observable. Thus, impractically large gains may be required to place all the poles from a single controller. Second, it is unclear that the approach uses the designer's available degrees of freedom in the best way. Essentially, the approach seems to require that all the disturbances in the system propagate to a single output, where they can be observed and compensated for by control signals at an adjacent input. In many situations, it would seem better to have controllers act to suppress the spread of disturbances through the system. Finally, concentration of all the complexity of the control structure at a single (or few) controllers may be undesirable.

The Davison–Wang approach suffers similar drawbacks. Although there is no explicit attempt to make all of the strongly connected subsystems controllable and observable from a single controller, the generic outcome of the first step of their approach will be precisely this situation.

It is interesting to compare the pole-placement and linear-quadratic-Gaussian approaches to the design of decentralized controllers. One pertinent observation is that, heuristically speaking at least, the signaling phenomena is basic to both approaches. We have already discussed signaling in the LQG context, so let us consider its role in the pole-placement problem. Consider for concreteness the example in Fig. 7, and in particular the role of the first controller (Fig. 7(c)). Closing the feedback loop from y_1 to u_1 has the effect of making x_2 observable from y_2, and thus, in a very concrete sense, signaling the value of $x_2(t)$ to the second controller. Note, however, that closing the loop from y_1 to u_1 also has the effect of making x_1 controllable from y_1. As we commented earlier, it seems impossible to ascertain precisely, and perhaps even meaningless to ask, what portion of a control law is used for signaling and what portion for control.

The role of signaling in qualitative (as opposed to optimal) decentralized control theory appears even more clearly in a recent paper of Kobyashi *et al.* [124]. The problem considered is the determination of conditions under which an arbitrary unknown initial condition of

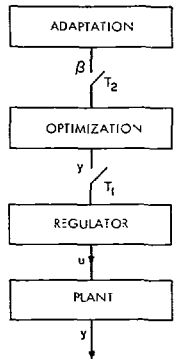

Fig. 8. Multilayer controller.

(4.13) can be driven to zero in finite time by decentralized control. The approach is to let the system run free for a short period of time during which each controller uses its observation $y_i(t)$ to learn the projection $P_i x(0)$ of the initial state $x(0)$ on the observable subspace corresponding to $[A, C_i]$. The controllers then (one at a time) signal the value of $P_i x(0)$ to the other controllers. After each controller finishes signaling, it removes the effect of its control on the system. Thus, under appropriate conditions, each controller learns the initial state $x(0)$ so that the controllers can cooperate in an open-loop fashion to drive the state to zero.

As a final comment on the results surveyed in this section, we note that the basic issue of what decentralized structure is desirable is left completely unresolved. What would be useful would be a simple characterization of, say, the class of decentralized structures that could stabilize a given system. But, of course, such a characterization is not available even in the centralized case.

V. Multilevel Methods for Deterministic Optimal Control

A. Introduction

Large industrial complexes consist of interconnected subsystems and in practice these are not controlled in a centralized manner. In trying to conceptualize this practice two types of control structure may be usefully distinguished. In a *multilayer* structure the determination of control is split into algorithms which operate at different time scales. A symbolic representation of a three layer structure is given in Fig. 8, [125], where the plant control variable u is determined by a regulator so as to bring specified output variables y to their desired target values y^d within a period T_1. The values y^d are specified by the optimization layer for a period $T_2 \gg T_1$. In computing these values the optimization layer assumes as given, during the period T_2, some environmental parameters θ. The adaptation layer specifies updated values of θ at the end of each period of length T_2. It is evident that such a structure does represent some aspects of the manner in which an industrial complex is controlled. Equally evident is the assertion that the higher layer operates using a more aggregated model than a lower layer. Unfortunately the literature dealing with the analysis or design of multilayer structures which goes beyond such descriptive assertions is scanty, although some of the papers dealing with singularly perturbed systems discussed in Section II may be considered as relevant. Here we will be concerned with the extensive literature on *multilevel* structures in which the control vector u is partitioned into subvectors each of which is independently chosen by separate local controller. A local controller's choice is based on its own simplified system model and simplified cost criterion. In order that these independently arrived at choices be coherent with the overall system criterion, a higher level controller is designed whose function is to coordinate the lower level controllers by manipulating their individual models or criteria. We shall illustrate the specific assumptions on the system model and the steps necessary to design a two-level controller.

B. Two-level Control Structure

Step 1: The overall system

$$\dot{x} = F(x,u,t) \tag{5.1}$$

is assumed to consist of isolated subsystems

$$\dot{z}_i = f_i(z_i,u_i,t) + g_i(v,t), \qquad i = 1,\cdots,k \tag{5.2}$$

interconnected by the constraint

$$v(t) = x(t). \tag{5.3}$$

Here $x \in R^n$, $z_i \in R^{n_i}$, $u \in R^m$, $u_i \in R^{m_i}$ and under (5.1), (5.2), and (5.3)

$$x'(t) = (z_1'(t),\cdots,z_k'(t)), \quad u'(t) = (u_1'(t),\cdots,u_k'(t)). \tag{5.4}$$

The choice of control is restricted by the vector inequality

$$u(t) \in U(x(t),t) = \{u | q(x(t),u,t) \leqslant 0\} \tag{5.5}$$

and the initial state is specified as $x(t_0) = x_0$. The objective of control is

$$\underset{u(t)\in U(x(t),t)}{\text{minimize}} \int_{t_0}^{t_1} c(x(t),u(t),t)\,dt. \tag{5.6}$$

All of the techniques require some form of *separability* which means that one or more of the following relations hold:

$$g_i(v,t) = \sum_j g_{ij}(v_j,t), \tag{5.7}$$

$$q(z,u,t) = \sum_j q_j(z_j,u_j,t) \tag{5.8}$$

$$c(z,u,t) = \sum_j c_j(z_j,u_j,t). \tag{5.9}$$

(In many cases a clever choice of additional variables can yield separability when at first glance this may not appear

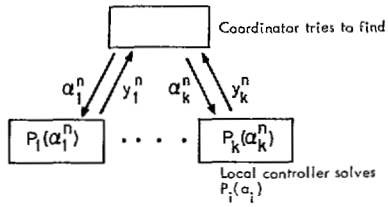

Fig. 9. Two-level controller.

possible.) Assuming that all of these relations hold, the original optimal control problem (5.1), (5.5), (5.6) can be rewritten as

$$\begin{aligned} &\text{minimize} \sum_i \int_{t_0}^{t_1} c_i(z_i, u_i, t)\, dt \\ &\text{subject to } \dot{z}_i = f_i(z_i, u_i, t) + \xi_i(t), \qquad i = 1, \cdots, k \end{aligned} \tag{5.10}$$

$$\left.\begin{aligned} &z_i(0) = x_{0i} \\ &\xi_i(t) = \sum_j g_{ij}(z_j, t), \qquad i = 1, \cdots, k \\ &\sum_i q_i(z_i, u_i, t) \leqslant \mathbf{0}. \end{aligned}\right\} \tag{5.11}$$

In summary, in Step 1 the original problem is transformed into a set of subproblems by the use of *interaction* variables $\xi_i(t)$. The fact that the original problem is not a collection of independent subproblems is reflected in the *coupling* constraints of (5.11). This reformulation of the problem is referred to as the *global* problem P_G [126]. Often the interconnected control constraint of (5.11) is assumed to be decoupled [127]

$$q_i(z_i, u_i, t) \leqslant 0, \qquad i = 1, \cdots, k, \tag{5.12}$$

and for simplicity this will be assumed below.

Step 2: (Decomposition). This is the critical step. The global problem is now replaced by a *parametrized family* of k decoupled subproblems of smaller size and simpler structure. Call the parameter $\alpha = (\alpha_1, \cdots, \alpha_k)$ and the subproblems $P_1(\alpha_1), \cdots, P_k(\alpha_k)$. This family must be such that P_G is *imbedded* in it, that is, there is a parameter α^* at which the solution to $P_i(\alpha_i^*)$, $i = 1, \cdots, k$, yields a solution to P_G. In the language of multilevel control (see [128], [129]) this imbedding condition is called *coordinability*, whereas if we adopt the terminology of the large-scale mathematical programming literature, it would be more appropriate to call the family of parametrized problems as the *master problem* [130]. The parameter α^* is initially unknown, and all techniques propose an iterative procedure of the following kind (see Fig. 9):

Step A: Local controller i receives α^n from the coordinator, solves $P_i(\alpha^n)$, and then transmits some function y_i^n of its solution to the coordinator.

Step B: The coordinator computes a new value $\alpha^{n+1} = (\alpha_1^{n+1}, \cdots, \alpha_k^{n+1})$ based on α^n and $y_1^n, \cdots, y_k^n$. Return to Step A.

C. Examples of Iterative Procedure

The simplification in the structure and size of the decoupled problems is invariably achieved by eliminating explicit treatment of the coupling constraint of (5.11).[15] We sketch the two basic ways of doing this.

In the first approach ξ_i is made a part of the parameter α_i, so that the ith local controller treats the interaction variable as given and ignores the coupling constraint. If we take $\alpha_i(\cdot) = \xi_i(\cdot)$, then the subproblem $P_i(\alpha_i)$ becomes

$$\text{minimize} \int_{t_0}^{t_1} c_i(z_i, u_i, t)\, dt \tag{5.13}$$

$$\left.\begin{aligned} \text{subject to} \quad &\dot{z}_i = f_i(z_i, u_i, t) + \alpha_i(t), \quad z_i(0) = x_{0i} \\ &q_i(z_i, u_i, t) \leqslant \mathbf{0} \end{aligned}\right\}. \tag{5.14}$$

Let $z_i(\alpha_i)$ be the optimal trajectory of $P_i(\alpha_i)$. Then the imbedding condition requires that the "error"

$$e_i(\alpha)(t) = \alpha_i(t) - \sum_j g_{ij}(z_j(\alpha), t) \tag{5.15}$$

vanish for some α^*. Step B of the procedure must be such that α^{n+1} can be computed knowing α^n, and so that $e(\alpha^n)$ converges to zero. For further specification and discussion of this approach see the "interaction prediction principle" of Mesarovic, Macko, and Takahara [128]–[129], [132], [135]–[137], and [152].

In the second approach ξ_i is considered part of the control variable available to the ith controller. The coordinator modifies the cost function of each local controller by dualizing the coupling constraint (5.10) (see [138, ch. 7] or [139]), giving the following problem in which α is a parameter:

$$\begin{aligned} \text{minimize} \quad &\sum_i \int_{t_0}^{t_1} c_i(z_i, u_i, t)\, dt + \sum_i \int_{t_0}^{t_1} \alpha_i^T(t) \\ &\cdot \left[\xi_i(t) - \sum_j g_{ij}(z_j, t)\right] dt \end{aligned} \tag{5.16}$$

$$\left.\begin{aligned} \text{subject to} \quad &\dot{z}_i = f_i(z_i, u_i, t) + \xi_i(t), && i = 1, \cdots, k \\ &z_i(0) = x_{0i}, && i = 1, \cdots, k \\ &q_i(z_i, u_i, t) \leqslant \mathbf{0}, \xi_i \text{ arbitrary}, && i = 1, \cdots, k \end{aligned}\right\}. \tag{5.17}$$

This evidently decomposes into k independent subproblems, $P_i(\alpha)$, when we observe that the cost function can be rearranged as

$$\sum_i \int_{t_0}^{t_1} \left[c_i(z_i, u_i, t) + \alpha_i^T \xi_i - \sum_j \alpha_j g_{ji}(z_i, t)\right] dt. \tag{5.18}$$

[15] The only exceptions to this are the "feasible" method of Grateloup and Titli [131], also called "model coordination" in [132]. For control problems these methods are usually inapplicable since they require that the dimension of u_i exceed the dimension of the interacting variables. These methods may be compared with the "primal" approaches in mathematical programming (see [133] and [134].

Let $z_i(\alpha)$, $u_i(\alpha)$, $\xi_i(\alpha)$ be the solution of $P_i(\alpha)$, and let $e_i(\alpha)(t)=\xi_i(\alpha)(t)-\sum_j g_{ij}(z_j(\alpha), t)$. The imbedding condition requires that there exist α^* for which $e(\alpha^*)=0$, which is difficult without some kind of convexity restriction. Once again Step B of the procedure must be such that α^{n+1} can be computed knowing α^n and $e(\alpha^n)$ so that $e(\alpha^{n+1})$ is smaller than $e(\alpha^n)$. A variety of methods have been proposed for doing this. For these and some applications see [134] and [140]–[143], the "interaction balance principle" of Mesarovic, Macko, and Takahara [129], and [144]–[152].

D. Evaluation of Multilevel Techniques

We evaluate the literature along several dimensions.

Computation. Adoption of a multilevel technique entails the replacement of a single large problem P_G by a sequence of subproblems $P_1,\cdots,P_k$ and a problem of coordination. Since computation time usually rises at a greater than linear rate with problem size it is generally true that solution of $P_1,\cdots,P_k$ is less costly than solution of P_G. However, the need to iterate and coordinate implies that any computational savings associated with multilevel techniques are problem dependent. It is difficult to make any universally valid general statements, but one situation in which decomposition definitely appears useful occurs when the P_i are not only smaller but also *simpler* than P_G. An example of this type is the Cantor–Gerla algorithm [154], in which P_i is solved by an efficient algorithm that is not applicable to P_G.

An advantage of decomposition algorithms is the ability to solve the P_i independently one from another. Thus if the global problem P_G has a primary memory requirement in excess of that available but there is sufficient memory for each P_i, the multilayer technique may be applicable even if P_G cannot be solved directly. Alternatively, there is the possibility of parallel solution of the P_i if multiple processors are available. However, we are unaware of any computational experience supporting these claims. Note that paging in modern virtual memory systems provides a solution to the memory problem that is often implemented in hardware and may thus be superior. Furthermore, perhaps the most likely architecture of future parallel processing computers is the so-called vector machine for which a decomposition algorithm would seem to present no particular advantage.

Information requirements. In both the examples given above, and in the cited literature, it is evident that neither the coordinator nor the local controller need to know the global problem, in particular the overall system model. Thus, *knowledge of the model is decentralized*, in comparison with a centralized or direct approach. The advantage appears overwhelming when k is very large and especially when the model of each isolated system is "internal" to the ith controller (this is clear when the controller is a person, as in the case of an economy, who is unable to communicate this model [140], [153]). But even this advantage is attenuated when we observe *first* that multilevel techniques are limited to *off-line calculation of open-loop control* (with some exceptions noted in the next section), and *second* that the information exchanged between the coordinator and the lower level controllers is of the same order of magnitude as would be necessary to communicate the entire state trajectory. Thus, if feedback is essential then these techniques become inapplicable, whereas if it is not, then the centralized knowledge of the state trajectory tends to reduce the advantage of the decentralized knowledge of the system model, especially if (t_1-t_0) is large.

Other considerations. It is sometimes alleged [126], [128], that multilevel optimization techniques are insensitive to modeling errors and component failures, but the arguments attempting to show this remain unconvincing. (The fact that hierarchies abound in nature and in human organizations is irrelevant.) Finally, except for the techniques of [131] and [132], the intermediate solutions produced in the iterative procedure are all unfeasible. This may be a nuisance.

E. Miscellaneous

It is possible to compute a *centralized* feedback law using an *off-line* multilevel technique. Once this has been done, the implementation—in particular, all *on-line* processing—is centralized. This has been shown for the linear, quadratic problem by Singh, Hassan, and Titli [152]. (However this should be sharply distinguished from those control techniques which seek to decentralize on-line processing as in Section IV.)

The multilevel optimal control techniques discussed in this section can be viewed as function space generalizations of corresponding decomposition algorithms for mathematical programming problems. This, in part, explains the emphasis on *open-loop and deterministic problems*. However, there is another decomposition theoretic literature that has developed in the context of solution of large, sparse sets of linear equations (see, e.g., [155]). These techniques can also be generalized to optimal control problems, [105], [156]. An advantage of this approach is that closed-loop and stochastic problems can be considered for which algorithms producing feasible intermediate solutions can be derived.

VI. Summary and Conclusions

We have divided the work surveyed into the four categories of Model Simplification, Stability of Interconnected Systems, Decentralized Control, and Hierarchical Control. This seems to us a fair categorization of at least a substantial portion of the literature on large scale systems. However, the problems of designing control systems for large scale systems do not so neatly subdivide. We feel that a fruitful area of research is the unification of the aforementioned research areas. To a certain extent, this unification

is already proceeding, and we hope that this paper speeds the process.

With respect to designing decentralized controllers for many physical large scale systems we believe that a good combination of engineering judgment and analysis can be used to define in a reasonable, albeit ad hoc, way a special structure for the dynamic system; one should look for time scale separation (fast and slow dynamics), weak coupling and similar phenomena, since they naturally lead to ways of decentralizing the system. The problem, as we have remarked, is to select the appropriate state space representation so that the weak or strong coupling parameters appear in the mathematical model of the system. Once the decentralized structure of the compensators has been defined, then the parameters of the compensators can be optimized.

Another method that may help in the selection of a decentralized structure, is to actually solve the centralized problem. In the case of LQG designs, one has to have *reliable* algorithms that solve very high-dimensional Riccati equations. One can then examine the relative magnitudes of the various Kalman filter gains and control gains, and eliminate those that are small. In this manner one may arrive at a decentralized estimation and control structure, and then reoptimize all of the filter and control gains to further improve the performance. There are two problems with this approach: first, reliable software is not readily available and second, for a sufficiently large system the sheer numbers of filter and control gains make the human judgmental task of isolating the important coupling terms extremely difficult. What is needed is a computer-aided procedure which presents to the designer several good candidates for further consideration, and which weeds out overly complex configurations. Such systematic procedures are not available.

Our most fundamental conclusion, after surveying a vast amount of literature, is that although considerable progress has been made in many directions, the question of what *structures* are desirable for control of large scale systems has not been addressed in a truly scientific fashion. In our opinion, we do not believe that the existing mathematical tools (such as dynamic programming) are powerful enough to define a preferable structure for decentralized and/or hierarchical control. First, we do not believe that it is reasonable to seek a *single best* optimal structure. Rather, any future methodologies should strive to define sets of distributed information and control structures that are preferable to others. Second, a unified theory of decentralized control should explicity include not only a traditional index of performance for the control system but in addition

a) the cost of communications
b) reliability issues
c) the cost of computer interfaces
d) the value of incomplete and/or delayed information
e) a formal measure of system complexity.

To develop such a desirable methodology we may have to develop different definitions of optimality, principles of optimality, and notions of optimal solutions.

ACKNOWLEDGMENT

Many colleagues have helped us with references, criticisms and suggestions. In particular we would like to acknowledge our gratitude to Professor A. N. Michel, to Professor Ted Davison, Dr. D. Castanon, Dr. A. Laub, and to the following MIT graduate students: J. D. Birdwell, D. Looze, J. Wall, and K. Yared. We also wish to thank several anonymous reviewers for their comments.

REFERENCES

[1] A. Gelb, Ed., *Applied Optimal Estimation*. Cambridge, MA: M.I.T. Press, 1974.
[2] M. Aoki, "Control of large-scale dynamic systems by aggregation," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 246–253, June 1968.
[3] B. P. Ziegler, "Towards a formal theory of modeling and simulation: Structure preserving morphisms," *J. Ass. Comput. Mach.*, vol. 19, no. 4, Oct. 1972.
[4] R. A. Howard, *Dynamic Probabilistic Systems*. New York: Wiley, 1971.
[5] J. S. Chipman, "Estimation and aggregation in econometrics: an application of the theory of generalized inverses," in *Generalized Inverses and Applications*, M. Z. Nasled, Ed. New York: Academic, 1976.
[6] L. Aldermeshian, "Model reduction techniques in the stochastic control of data communication networks," M.S. thesis, Mass. Inst. Tech., Cambridge, 1975
[7] E. J. Davison, "A method for simplifying linear dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 93–101, Jan. 1966.
[8] R. M. Chidamabara, "On 'A method for simplifying linear dynamic systems,'" *IEEE Trans. Automat. Contr.*, vol. AC-12, pp. 119–121, Feb. 1967.
[9] E. J. Davison, "A new method for simplifying linear dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-13, pp. 214–215, Apr. 1968.
[10] M. R. Chidamabara, "Two simple techniques for the simplification of large dynamic systems," in *Proc. Joint Automatic Control Conf.*, Univ. Colorado, Boulder, Aug. 1969.
[11] D. A. Wilson, "Optimum solution of model reduction problem," *Proc. IEE*, vol. 117, June 1970.
[12] J. M. Siret, G. Michailesco, and P. Bertrand, "Representation of linear dynamical systems by aggregated models," to be published.
[13] J. E. Van Ness, H. Zimmer, and M. Cultu, "Reduction of dynamic models of power systems," *1973 PICA Proc.*, Minneapolis, MN, June 1973.
[14] M. Aoki, "Note on aggregation and bounds for the solution of the matrix Riccati equations," *J. Math. Anal. Appl.*, vol. 21, 1968.
[15] R. B. Schainker, "Suboptimal control by aggregation," *Proc. Joint Automatic Control Conf.*, Atlanta, GA, June 1970.
[16] H. J. Kwatny, J. H. Spare, F. M. Massimo, and A. D. Bhatt, "Perturbation methods in the construction of model decomposition for large scale systems analysis," Drexel Univ., Rep., Aug. 1977.
[17] H. K. Ramapriyan, "Study of coupling in interconnected systems," Ph.D. dissertation, Univ. Minnesota, Minneapolis, Mar. 1970.
[18] F. N. Bailey and H. K. Ramapriyan, "Bounds on suboptimality in the control of linear dynamic systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 532–534, Oct. 1973.
[19] D. Teneketzis, "Perturbation methods in decentralized stochastic control," M.S. thesis, Mass. Inst. Tech., Cambridge, June 1976.
[20] D. Teneketzis, and N. R. Sandell, Jr., "Perturbation methods in decentralized stochastic control," *1st Int. Conf. on Information Sciences and Systems*, Patras, Greece, Aug. 1976.
[21] P. V. Kokotovic, "Feedback design of large linear systems," in *Feedback Systems*, J. B. Cruz, Ed. New York: McGraw-Hill, 1972.
[22] R. H. Kwong, "On a stochastic control method for weakly coupled linear systems," M.S. thesis, Mass. Inst. Tech., Cambridge, 1972.
[23] T. M. Athay and N. R. Sandell, Jr., "An iterative decoupling

method for large scale Lyapunov equations," *8th Ann. Southeastern Symp. on Systems Theory*, Knoxville, TN, Apr. 1976.
[24] N. R. Sandell, Jr., "Decomposition vs. decentralization in large scale system theory," *1976 IEEE Conf. on Decision and Control*, Clearwater, FL, Dec. 1976.
[25] P. V. Kokotovic, "Separation of time scales in modeling and control," *1975 IEEE Conf. on Decision and Control*, Houston, TX, Dec. 1975.
[26] R. D. Milne, "The analysis of weakly coupled dynamical systems," *Int. J. Contr.*, vol. 2, pp. 171–199, Aug. 1965.
[27] P. V. Kokotovic and A. H. Haddad, "Controllability and time-optimal control of systems with slow and fast modes," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 111–113, Feb. 1975.
[28] P. V. Kokotovic and R. A. Yackel, "Singular perturbation of linear regulators: basic theorems," *IEEE Trans. Automat. Contr.*, vol. AC-17, 29–37, Feb. 1972.
[29] R. A. Yackel and P. V. Kokotovic, "A boundary layer method for the matrix Riccati equations," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 17–24, Feb. 1973.
[30] C. A. Desoer and M. J. Shensa, "Networks with very small and very large parasitics: natural frequencies and stability," Electronics Research Lab., Memo ERL-M276, Univ. of Calif., Berkeley, Apr. 1970.
[31] A. H. Haddad and P. V. Kokotovic, "On singular perturbations in linear filtering and smoothing," *Proc. 5th Symp. Nonlinear Estimation and Applications*, Sept. 1974.
[32] A. H. Haddad, "Linear filtering of singularly perturbed systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 515–519, Aug. 1976.
[33] D. Teneketzis and N. R. Sandell, Jr., "Linear regulator design for stochastic systems by a multiple time scales method," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 615–621, Aug. 1977.
[34] A. H. Haddad and P. V. Kokotovic, "Stochastic control of linear singularly perturbed systems," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 815–821, Oct. 1976.
[35] U. Ozguner, "Near optimal control of composite systems: the multi-time-scale approach," to be published.
[36] J. H. Chow and P. V. Kokotovic, "Eigenvalue placement in two-time-scale systems," in *Large Scale Systems Theory and Applications*. G. Guardabassi and A. Locatelli, Ed., IFAC Symp., Udine, Italy, June 1976.
[37] M. Ikedo and S. Kodema, "Large-scale dynamical systems: state equations, Lipschitz conditions and linearization," *IEEE Trans. Circuit Theory*, vol. CT-20, 1973.
[38] S. P. Singh and R. W. Liu, "Existence of state equation representation of linear large-scale dynamical systems," *IEEE Trans. Circuit Theory*, vol. CT-20, 1973.
[39] H. H. Rosenbrock and A. C. Pugh, "Contributions to a hierarchical theory of systems," *Int. J. Contr.*, vol. 19, no. 5, 1974.
[40] M. L. J. Hautus, "Input regularity of cascaded systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 120–123, Feb. 1975.
[41] E. J. Davison and S. H. Wang, "New results on the controllability and observability of general composite systems," *IEEE Trans. Automat. Contr.*, vol. AC-20, pp. 123–128, Feb. 1975.
[42] W. Hahn, *Theory and Application of Lyapunov's Direct Method*. Englewood Cliffs, NJ: Prentice-Hall, 1963.
[43] C. A. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Properties*. New York: Academic, 1975.
[44] V. Lakshmikantham and S. Leela, *Differential and Integral Inequalities*, vol. 1. New York: Academic, 1969.
[45] F. N. Bailey, "The application of Lyapunov's second method to interconnected systems," *SIAM J. Contr.*, vol. 3, no. 3, 1966.
[46] S. M. Cuk and D. D. Siljak, "Decomposition-aggregation stability analysis of the spinning skylab," *Proc. 7th Asilomar Conf. on Circuits, Systems, and Computers*, Pacific Grove, CA, 1973.
[47] V. M. Matrosov, "Method of vector Lyapunov functions in feedback systems," *Automat. Remote Contr.*, vol. 33, no. 9, Sept. 1972.
[48] A. A. Piontkovskii and L. D. Rutkovskaya, "Investigation of certain stability-theory problems by the vector Lyapunov function method," *Automat. Remote Contr.*, vol. 10, Oct. 1967.
[49] D. D. Siljak, "Stability of large-scale systems," in *Proc. 5th IFAC World Cong.*, Paris, France, June 1972.
[50] ——, "Large-scale systems: complexity, stability, reliability," in *Proc. Utah State Univ., Ames Research Center Seminar Workshop on Large Scale Dynamic Systems*, Utah State Univ., Logan, Aug. 1974.
[51] ——, "On the connective stability and instability of competitive equilibrium," in *Proc. Joint Automatic Control Conf.*, Univ. Texas, Austin, 1974.
[52] ——, "When is a complex ecosystem stable?," *Mathematical Biosciences*, vol. 25, 1975.
[53] ——, "Competitive economic systems: stability, decomposition, and aggregation," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 149–160, Apr. 1976.
[54] M. Araki and B. Kondo, "Stability and transient behavior of composite nonlinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 537–541, Aug. 1972.
[55] M. Araki, K. Ando, and B. Kondo, "Stability of sampled-data composite systems with many nonlinearities," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 22–27, Feb. 1971.
[56] R. F. Estrada, "On the stability of multiloop feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 781–791, Dec. 1972.
[57] L. T. Grujic and D. D. Siljak, "Asymptotic stability and instability of large-scale systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 636–645, Dec. 1973.
[58] W. E. Thompson, "Exponential stability of interconnected systems," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 504–506, Aug. 1970.
[59] A. N. Michel, "Stability analysis of interconnected systems," *SIAM J. Contr.*, vol. 12, Aug. 1974.
[60] L. T. Grujic and D. D. Siljak, "On the stability of discrete composite systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 522–524, Oct. 1973.
[61] F. N. Bailey, "The concept of aggregation in system stability analysis," *Asilomar Conf. on Circuits, Systems, and Computers*, Asilomar, CA, Nov. 1968.
[62] M. A. Pai and C. L. Narayana, "Stability of large scale power systems," *Proc. IFAC World Cong.*, Boston, MA, Aug. 1975.
[63] G. Zames, "On the input–output stability of time-varying nonlinear feedback systems: Part I," *IEEE Trans. Automat. Contr.*, vol. AC-11, pp. 228–238, Apr. 1966.
[64] E. J. Lasley and A. N. Michel, "Input–output stability of interconnected systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 84–89, Feb. 1976.
[65] P. A. Cook, "On the stability of interconnected systems," *Int. J. Contr.*, vol. 20, no. 3, 1974.
[66] D. W. Porter and A. N. Michel, "Input-output stability of time-varying nonlinear multiloop feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 422–427, Aug. 1974.
[67] M. Araki, "Input–output stability of composite feedback systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 254–259, Apr. 1976.
[68] P. A. Cook, "Modified multivariable circle theorems," in *Recent Mathematical Developments in Control*, D. J. Bell, Ed. New York: Academic, 1973.
[69] H. H. Rosenbrock, "Progress in the Design of Multivariable Control Systems," *ASME Trans. Inst. Measure. Contr.*, vol. 4, 1971.
[70] ——, "Multivariable circle theorems," in *Recent Mathematical Developments in Control*, D. J. Bell, Ed. New York: Academic, 1973.
[71] M. K. Sundareshan and M. Vidyasagar, "L_2-stability of large-scale dynamical systems—criteria via positive operator theory," to be published.
[72] J. C. Willems, "Qualitative behavior of interconnected systems," *Annals of Syst. Res.*, vol. 3, 1973.
[73] M. G. Safonov, "Robustness and stability aspects of stochastic multivariable feedback system design," Ph.D. dissertation Mass. Inst. Tech., Sept. 1977.
[74] F. M. Callier, W. S. Chan, and C. A. Desoer, "Input-output stability theory of interconnected systems using decomposition techniques," *IEEE Trans. Circuits Syst.*, vol. CAS-23, Dec. 1976.
[75] A. N. Michel and R. D. Rasmussen, "Stability of stochastic composite systems," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 89–94, Feb. 1976.
[76] R. R. Rasmussen and A. N. Michel, "On vector Lyapunov functions for stochastic dynamical systems," *IEEE Trans. Automat Contr.*, vol. AC-21, pp. 250–254, Apr. 1976.
[77] A. N. Michel, "Quantitative analysis of simple interconnected systems: stability boundedness, and trajectory behavior," *IEEE Trans. Circuit Theory*, vol. CT-17, Aug. 1970.
[78] ——, "Stability and trajectory behavior of composite systems," *IEEE Trans. Circuits Syst.*, vol. CAS-22, pp. 305–312, Apr. 1975.
[79] L. T. Grujic, "Vector Lyapunov functions and singularly perturbed large-scale systems," in *Proc. 1976 Joint Automat. Contr. Conf.*, Purdue Univ., LaFayette, IN, 1976.
[80] A. N. Michel and R. K. Miller, *Qualitative Analysis of Large Scale Dynamical Systems*. New York: Academic, 1977.
[81] H. Kwakernaak and R. Sivan, "The maximally achievable accuracy of linear optimal regulators and linear optimal filters," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 79–86, Feb. 1972.
[82] M. G. Safonov and M. Athans, "Gain and phase margins for multiloop LQG regulators," *IEEE Trans. Automat. Contr.*, vol. AC-22, pp. 173–177, Apr. 1977.
[83] H. S. Witsenhausen, "Separation of estimation and control for discrete time systems," *Proc. IEEE*, vol. 59, Nov. 1971.
[84] V. E. Benes, "Existence of optimal stochastic control laws," *SIAM J. Contr.*, vol. 9, Aug. 1971.
[85] H. S. Witsenhausen, "A counterexample in stochastic optimal control," *SIAM J. Contr.*, vol. 6, no. 1, 1968.
[86] M. Aoki, "On decentralized linear stochastic control problems

with quadratic cost," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 243–250, June 1973.
[87] J. M. Bismut, "An example of interaction between information and control: The transparency of a game," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 518–522, Oct. 1973.
[88] N. R. Sandell, Jr. and M. Athans, "Solution of some nonclassical LQG stochastic decision problems," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 108–116, Apr. 1974.
[89] ——, "A finite state, finite-memory minimum principle," *J. Optimiz. Theory Appl.*, vol. 25, no. 2, 1978.
[90] Y. C. Ho and K. C. Chu, "Team decision theory and information structures in optimal control problems, Part I," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 15–22, Feb. 1972.
[91] K. C. Chu, "Team decision theory and information structures in optimal control problems, Part II," *IEEE Trans. Automat. Contr.*, vol. AC-17, pp. 22–28, Feb. 1972.
[92] Y. C. Ho and K. C. Chu, "Information structure in dynamic multi-person control problems," *Automatica*, vol. 10, July 1974.
[93] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1965.
[94] Y. C. Ho, M. P. Kastner, and E. Wong, "Teams, signaling, and information theory," this issue, pp. 305–312.
[95] H. S. Witsenhausen, "A standard form for sequential stochastic control," *J. Math. System Theory*, vol. 7, no. 1, 1973.
[96] I. B. Rhodes and D. G. Luenberger, "Stochastic differential games with constrained state estimators," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 476–481, Oct. 1969.
[97] W. Willman, "Formal solutions for a class of stochastic pursuit–evasion games," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 504–509, Oct. 1969.
[98] P. Whittle and J. Rudge, "The optimal linear solution of a symmetric team control problem," *J. Appl. Probability*, vol. 11, 1974.
[99] C. Y. Chong and M. Athans, "On the stochastic control of linear systems with different information sets," *IEEE Trans. Automat. Contr.*, vol. AC-16, pp. 423–430, Oct. 1971.
[100] R. L. Kosut, "Suboptimal control of linear time-invariant systems subject to control structure constraints," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 557–563, Oct. 1970.
[101] F. C. H. Wang, "Suboptimization of decentralized control systems," Ph.D. dissertation, Univ. of Minnesota, Minneapolis, 1972.
[102] N. R. Sandell, Jr., "Information flow in decentralized stochastic systems," in *Directions in large Scale Systems*, Y. C. Ho and S. K. Mitter, Ed. New York: Plenum Press, 1976.
[103] C. W. Sanders, E. C. Tacker, and T. A. Linton, "A new class of decentralized filters for interconnected systems," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 259–262, June 1974.
[104] T. D. Linton, E. C. Tacker, C. W. Sanders, and T-C. Wang, "Computational and performance aspects of static decentralized controllers," in *Proc. 1976 IEEE Conf. on Decision and Control*, Clearwater, FL, Dec. 1976.
[105] D. P. Looze and N. R. Sandell, Jr., "Decentralization and decomposition in linear stochastic control," in *Proc. 1977 Joint Automatic Control Conf.*, San Francisco, CA, June 1977.
[106] ——, "Decomposition of linear, decentralized stochastic control Problems," *IFAC Workshop on Control and Management of Integrated Industrial Complexes*, Toulouse, France, Sept. 1977.
[107] K. C. Chu, "Optimal decentralized regulation for a string of coupled systems," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 243–246, June 1974.
[108] E. J. Davison, N. S. Rau, and F. V. Palmay, "The optimal decentralized control of a power system consisting of a number of interconnected synchronous machines," *Int. J. Contr.*, vol. 18, no. 6, 1973.
[109] D. P. Looze, P. K. Houpt, N. R. Sandell, Jr., and M. Athans, "On suboptimal decentralized estimation and control with application to freeway ramp monitoring," in *Proc. 1977 Joint Automatic Control Conf.*, San Francisco, CA, 1977.
[110] C. Y. Chong and M. Athans, "On the periodic coordination of linear stochastic systems," *Automatica*, vol. 12, July 1976.
[111] A. Benveniste, P. Bernhard, and G. Cohen, "On the decomposition of stochastic control problems," IRIA, France, Rapport de Recherche, no. 187, 1976.
[112] F. M. Brasch, Jr., and J. B. Pearson, "Pole placement using dynamic compensators," *IEEE Trans. Automat. Contr.*, vol. AC-15, pp. 34–43, Feb. 1970.
[113] D. McFadden, "On the controllability of decentralized microeconomic systems: the assignment problem," in *Mathematical Systems Theory and Applications*. New York: Springer-Verlag, 1967.
[114] M. Aoki, "On feedback stabilizability of decentralized dynamic systems," *Automatica*, vol. 8, no. 2, 1972.
[115] R. P. Corfmat and A. S. Morse, "Stabilization with decentralized feedback control." *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 679–682, Dec. 1973.
[116] M. Aoki and M-T. Li, "Controllability and stabilizability of decentralized dynamic systems," in *Proc. 1973 Joint Automatic Control Conf.*, Ohio State Univ., Columbus, July 1973.
[117] S. H. Wang and E. J. Davison, "On the stabilization of decentralized control systems," *IEEE Trans. Automat. Contr.*, vol. AC-18, pp. 473–478, Oct. 1973.
[118] J. P. Corfmat and A. S. Morse, "Control of linear system through specified input channels," *SIAM J. Contr.*, vol. 14, Jan. 1976.
[119] ——, "Decentralized control of linear multivariable systems," *Automatica*, vol. 12, Sept. 1976.
[120] E. J. Davison, "The robust decentralized control of a general servomechanism problem," *IEEE Trans. Automat. Contr.*, vol. AC-21, pp. 14–24, Feb. 1976.
[121] ——, "Decentralized stabilization and regulation," in *Directions in Large-Scale Systems*, Y. C. Ho and S. K. Mitter, Ed. New York: Plenum, 1976.
[122] A. S. Morse, "Structural invariants of linear multivariable systems," *SIAM J. Contr.*, vol. 11, Aug. 1973.
[123] E. J. Davison and S. H. Wang, "Properties and calculation of transmission zeros of linear multivariable systems," *Automatica*, vol. 10, 1974.
[124] H. Kobayashi, H. Hanafusu, and T. Yoshikawa, "Controllability of continuous-time decentralized control systems," *IEEE Trans. Automat. Contr.*, to be published.
[125] W. Findeisen, "A survey of problems in hierarchical control," *Workshop on Multilevel Control*, Inst. of Automat. Contr., Technical University of Warsaw, Warsaw, Poland, 1975.
[126] A. Benveniste, P. Bernhard, and A. Cohen, "On the decomposition of stochastic control problems," IRIA, France, Rapport de Recherche, no. 187, 1976.
[127] M. G. Singh, S. A. W. Drew, and J. F. Coales, "Comparisons of practical hierarchical control methods for interconnected dynamical systems," *Automatica*, vol. 11, 1975.
[128] M. D. Mesarovic, D. Macko, and Y. Takahara, *Theory of Hierarchical Multilevel Systems*. New York: Academic, 1970.
[129] ——, "Two coordination principles and their application in large scale systems control," *Automatica*, vol. 6, 1970.
[130] A. M. Geoffrion, "Elements of large-scale mathematical programming: Part I, concepts, Part II, synthesis of algorithms and bibliography," *Management Sci.*, vol. 16, 1970.
[131] G. Grateloup and A. Titli, "Two-level dynamic optimization methods," *J. Optimiz. Theory Appl.*, vol. 15, no. 3, 1975.
[132] J. D. Pearson, "Dynamic decomposition techniques," in *Optimization Methods for Large Scale Systems with Applications*, D. A. Wismer, Ed., New York: McGraw-Hill, 1971.
[133] G. B. Dantzig, "Linear control processes and mathematical programming," *SIAM J. Contr.*, vol. 4, no. 1, 1960.
[134] A. M. Geoffrion, "Primal resouce-directives approaches for optimizing nonlinear decomposable systems," *Oper. Res.*, vol. 18, no. 3, 1970.
[135] G. Cohen, "On an algorithm of decentralized optimal control," *J. Math. Anal. Appl.*, 1976.
[136] G. Cohen and G. Jolland, "Coordination methods by the prediction principle in large dynamic constrained optimization problems," in *Proc. IFAC Symp. LSSTA (Large Scale Systems Theory and Applications)*, Udine, Italy, 1976.
[137] G. Cohen, "A game theoretic approach to decentralized control," *Appl. Math.*, 1976.
[138] D. G. Luenberger, *Optimization by Vector Space Methods*. New York: Wiley, 1969.
[139] A. M. Geoffrion, "Duality in nonlinear programming: a simplified applications-oriented development," *SIAM Rev.*, vol. 13, no. 1, 1971.
[140] K. J. Arrow and L. Hurwicz, "Decentralization and computation in resource allocation processes," in *Mathematical Methods in the Social Sciences*, K. Arrow, S. Karlin and P. Suppes, Ed. Stanford, CA: Stanford Univ., 1960, pp. 27–46.
[141] L. M. Lasdon, "Duality and decomposition in mathematical programming," *IEEE Trans. Syst. Sci. and Cybern.*, vol. SSC-4, no. 2, 1968.
[142] H. Tamura, "Application of duality and decomposition in high order multi-stage decision processes," Cambridge Univ. Rep. no. CUED/B-Control/TR49.
[143] ——, "A discrete dynamic model with distributed transport delays and its hierarchical optimization for preserving stream quality," *IEEE Trans. Syst., Man and Cybern.*, vol. SMC-4, 1976.
[144] P. Bernhard, *Commande Optimale, Decentralisation et Jeux Dynamiques*. Paris, France: Dunod, 1976.
[145] L. S. Lasdon and J. D. Schoeffler, "A multilevel technique for optimization," in *Proc. Joint Automatic Control Conf.*, Troy, NY, 1965.
[146] M. G. Singh and H. Tamura, "Modeling and hierarchical optimization for oversaturated urban road traffic networks," *Int. J. Contr.*, vol. 20, 1974.
[147] A. Titli, T. Lefenie, and M. Richetin, "Multilevel optimization methods for the separable problems and applications," *Int. J. Syst. Sci.*, vol. 4, 1973.
[148] Y. Takahara and M. D. Mesarovic, "Coordinability of dynamic

systems," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 688–698, Dec. 1969.
[149] N. Smith and A. P. Sage, "An introduction to hierarchical systems theory," *Comput. Elec. Eng.*, vol. 1, 1973.
[150] F. Fallside and P. Perry, "Hierarchical optimization for a water supply network," in *Proc. IEEE*, vol. 122, no. 2, 1975.
[151] J. Bauman, "Multilevel optimisation techniques with application to trajectory decomposition," in *Advances in Control Systems*, C. T. Leondes, Ed. New York: Academic, 1968.
[152] M. G. Singh, M. F. Hassan, and A. Titli, "Multilevel feedback control for interconnected dynamical systems using the prediction principle," *IEEE Trans. Syst., Man and Cybern.* vol. SMC-6, no. 4, 1976.
[153] L. Hurwicz, "Optimality and informational efficiency in resource allocation processes," in *Mathematical Methods in the Social Sciences*, K. Arrow, S. Karlin and P. Suppes, Ed. Stanford CA: Stanford Univ., 1960, pp. 27–46.
[154] D. G. Cantor and M. Gerla, "Optimal routing in a packet-switched Computer network," *IEEE Trans. Comput.*, vol. C-23, Oct. 1974.
[155] R. S. Varga, *Matrix Iterative Analysis*. New York: Prentice-Hall, 1963.
[156] F. N. Bailey and A. J. Laub, "An iterative procedure of decentralized estimation and control," in *Proc. of the Workshop Discussion on Multilevel Control*, W. Findeisen, Ed. Warsaw, Poland, 1975.

Nils R. Sandell, Jr. (S'70–M'74) was born in Brooklyn, NY, on June 13, 1948. He received the B.E.E. degree from the University of Minnesota, Minneapolis, in 1970 and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1971 and 1974, respectively.

Since 1974 he has been on the faculty of the Electrical Engineering and Computer Science Department of the Massachusetts Institute of Technology, where he is presently an Associate Professor. He has held summer positions with Bell Telephone Laboratories (1969)–(1970), Honeywell Systems and Research Center (1972), the Massachusetts Institute of Technology Center for Advanced Engineering Study (1973), and The Analytic Sciences Corporation (1974). He has served as a consultant to The Analytic Sciences Corporation, Adcole Corporation, the Lincoln Laboratories of the Massachusetts Institute of Technology, and the Honeywell Systems and Research Center. In 1977 he received the Donald P. Eckman Award of the American Automatic Control Council.

Dr. Sandell is a member of Eta Kappa Nu, Tau Beta Pi, and Sigma Xi.

Pravin Varaiya was born in 1940 in Bombay, India.

He was a Guggenheim Fellow in 1971 and a Visiting Professor at M.I.T. during the academic year 1974–1975. He has also taught at COPPE, Federal University of Rio de Janeiro, Brazil. He is currently a Professor of Electrical Engineering and Computer Sciences, and of Economics, Univeristy of California, Berkeley. He is an Associate Editor of the *Journal of Optimization Theory and Applications*, and of *Revue Française d'Automatique, d'Informatique et de Recherche Opérationnelle* (RAIRO). He is author of *Notes on Optimization* (Van Nostrand Reinhold). He is actively engaged in teaching and research in stochastic control, large-scale systems, and Cuban economics.

Michael Athans (S'58–M'61–SM'69–F'73) was born in Drama, Greece on May 3, 1937. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1958, 1959, and 1961, respectively.

From 1961 to 1964 he was employed by the M.I.T. Lincoln Laboratory, Lexington, MA. In 1964, he joined the faculty in the Department of Electrical Engineering at the Massachusetts Institute of Technology where he currently has the rank of Professor and Director of the M.I.T. Electronics Systems Laboratory. He has been a consultant to the M.I.T. Lincoln Laboratory, Bell Aerosystems, Buffalo, NY, Bolt, Beranek and Newman, Cambridge, MA, Hamilton-Standard, Windsor Locks, CT, Systems Control, Inc., Palo Alto, CA, the U.S. Army Material Command, and the Analytic Sciences Corporation, Reading, MA His current research interests involve the theory and applications of optimal control and estimation techniques to aerospace, transportation, communications and socio-economic systems. He is the co-author of more than one hundred articles and co-author of the books *Optimal Control* (McGraw-Hill, 1966), *Systems, Networks and Computation*: Basic Concepts (McGraw-Hill, 1972), and *Systems, Networks and Computation: Multivariable Methods* (McGraw-Hill, 1974).

Dr. Athans is a member of AAAS, Phi Beta Kappa, Eta Kappa Nu and Sigma Xi. He was the recipient of the 1964 Donald P. Eckman Award. In 1969 he received the first Frederick Emmons Terman Award as the outstanding young electrical engineer educator, presented by the Electrical Engineering Division of the American Society for Engineering Education. He received honorable mention for a paper he co-authored for the 1971 JACC. He has been Program Chairman of the 1968 Joint Automatic Control Conference and member of several committees of AACC and IEEE. From 1972 to 1974 he was the President of the IEEE Control Systems Society. At present he is Associate Editor of the IFAC Journal *Automatica* and IEEE delegate to the American Automatic Control Council.

Michael G. Safonov (M'73–S'76–M'77) was born in Pasadena, CA, on November 1, 1948. He received the B.S., M.S., Engineer, and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, in 1971, 1971, 1976, and 1977, respectively.

From 1972 to 1975 he was on active duty with the U.S. Navy serving as electronics division officer abroad an aircraft carrier. He has held summer positions with the Air Force Cambridge Research Laboratories (AFCRL) (1968–1971) and has been a Consultant to The Analytic Sciences Corporation (1976). In 1977 he joined the faculty of the University of Southern California, Los Angeles, where he is presently an Assistant Professor of Electrical Engineering. His current research interests center on the fundamentals of input-output stability theory and on stability-based methods for the design of estimators and feedback controllers for nonlinear time-varying systems.

Dr. Safonov is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi.