ON THE USE OF QUASILINEARIZATION FOR THE SOLUTION OF SUB-PROBLEMS IN ON-LINE HIERARCHICAL CONTROL AND ITS APPLICATION TO A WATER DISTRIBUTION NETWORK

F. Fallside* and P.F. Perry**

Summary

Due to the difficulties of handling non-linearities in many large systems to which on-line optimal control is being applied, many applications have had to be restricted to the linear model-quadratic cost case. In particular, if the calculations are performed in a decentralized manner, the sub-system problems must yield a rapid solution and in the simple linear-quadratic case analytic solutions to these sub-problems may be obtained. The method of quasilinearization for the resolution of boundary-value problems arising in the solution of non-linear differential equations has been widely developed. This paper examines the use of quasilinearization algorithms for the solution of sub-problems arising in a problem decomposition using Lagrangian duality. The good convergence properties of the algorithms make them particularly useful for the solution of the sub-system problems. The actual improvement in operating costs obtained by handling more general sub-system non-linearities is compared with the increased computational burden for an actual on-line water control scheme.

* Reader in Electrical Engineering, Cambridge University.

** Research student in Electrical Engineering, Cambridge University.

1. Introduction

The determination of optimal controls for on-line implementation is, in general, a most difficult task. Only in restricted cases, such as linearized process models and quadratic cost expansions, can analytic solutions to the control problems be obtained, and consequently determination of optimal controls and the associated optimal trajectories must often be obtained iteratively on a computer. Since many practical systems are non-linear it is desirable to have a control algorithm capable of handling such non-linearities. The dynamic programming method[1-3] is capable of handling a very general type of problem, but has enormous high speed memory and computational requirements. State increment dynamic programming[4] reduces high speed memory requirements but still requires much computing time, while the successive approximation method[4] requires an equal number of state and control variables, although the time requirements are somewhat reduced. Thus a considerable amount of computing power has to be sacrificed if more general systems are to be handled using dynamic programming methods.

For the solution of very large problems formulated in a decentralized manner clearly none of the above methods are appropriate, since any hierarchical control algorithm involves the repetitive solution of many sub-system problems. The quasilinearization method,[5] developed by Bellman and Kalaba, is a useful method of obtaining an approximate solution to the non-linear sub-system boundary-value problem. It basically solves a sequence of linearized problems which hopefully converge on the true solution of the non-linear problem. Although convergence is in no sense guaranteed, in cases where it does work it provides a computationally efficient method of solving the non-linear sub-system problems.

If this method is to be of use in solving many non-linear sub-system problems in a hierarchical structure, clearly it must converge very rapidly. In this paper the quasilinearization algorithm is used for the solution of non-linear sub-system problems arising from the spatial decomposition technique used in Lagrange duality theory.[7] A modified[6] algorithm, which uses a scaling factor in the sub-system variations and has a descent property in the error indices, is assessed for its usefulness in speeding up sub-system convergence. The method is applied to a water supply control scheme, which includes sub-problems of a fairly general form, involving the minimization of a functional subject to differential, non-differential, inequality and terminal constraints. A comparison is made between pump energy requirements as predicted by linear and non-linear models at the sub-systems, and the increased computational burden of handling the non-linear case.

2. Statement of problem

The subject of this paper is the optimal control of a large dynamical system whose hierarchical structure allows it to be described as a set of N interconnected dynamical sub-problems[7] in the following way:

$$\min_{\underline{x}_i,\underline{u}_i,\underline{m}_i,\underline{y}_i} \sum_{i=1}^{N} \{F_i(\underline{x}_i(t_f)) + \int_{t_o}^{t_f} f_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t)\,dt\}$$

subject to

$$\begin{aligned}
\dot{\underline{x}}_i &= g_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t) \\
\underline{x}_i(t_o) &= \underline{x}_{io} \\
\underline{h}_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t) &\leq \underline{0} \\
\underline{H}_i(\underline{x}_i(t_f)) &\leq \underline{0} \\
\underline{q}_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t) &= \underline{0} \\
\underline{Q}_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t_f) &= \underline{0}
\end{aligned} \qquad i = 1,2,\ldots,N \qquad (1)$$

$$\begin{aligned}
\sum_{i=1}^{N} \underline{G}_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t) &= \underline{0} \\
\sum_{i=1}^{N} \underline{R}_i(\underline{x}_i,\underline{u}_i,\underline{y}_i,\underline{m}_i,t) &\leq \underline{0}
\end{aligned} \qquad (2)$$

where $\underline{x}_i$ are the states, $\underline{u}_i$ the controls, $\underline{y}_i$ the dependent variables, $\underline{m}_i$ the coordinating variables, F_i the terminal cost, f_i the instantaneous cost, all for the i^{th} sub-problem. Using the additive separability of the Lagrangian,[7] the following N sub-problems may be defined

$$\min_{\underline{x}_i,\underline{u}_i,\underline{m}_i} \{F_i(\underline{x}_i(t_f)) + \int_{t_o}^{t_f} f_i^* \,dt\}$$

subject to constraints (1) and (2), where

$$f_i^* = f_i + \langle\underline{\beta},\underline{G}_i\rangle + \langle\underline{\gamma},\underline{R}_i\rangle \qquad (3)$$

where $\underline{\beta}$ and $\underline{\gamma}$ are appropriate N-vector multipliers, giving the well-known Goal Coordination Algorithm which maximizes the dual function at one level and repetitively solves a set of N sub-problems at a lower level until overall convergence is reached. A similar decomposition in discrete time may also be obtained. This paper is primarily concerned with the derivation of efficient solution methods for general sub-problems of this form within this hierarchical structure. In particular, if true model non-linearities are used, is the increased computational burden associated with obtaining solutions by a modified quasilinearization method worthwhile in terms of the increased cost savings over the results for the linear case?

Ignoring the subscripts, clearly each sub-system problem is of the form

$$\min_{\underline{x},\underline{u},\underline{m},\underline{y}} J = F(\underline{x}(t_f)) + \int_{t_o}^{t_f} f(\underline{x},\underline{u},\underline{y},\underline{m},t)\,dt$$

where

$$\begin{aligned} \dot{\underline{x}} &= \underline{g}(\underline{x},\underline{u},\underline{y},\underline{m},t) \\ \underline{x}(t_o) &= \underline{x}_o \\ \underline{h}(\underline{x},\underline{u},\underline{y},\underline{m},t) &\leqslant \underline{0} \\ \underline{h}_f(\underline{x},\underline{u},\underline{y},\underline{m},t_f) &\leqslant \underline{0} \\ \underline{q}(\underline{x},\underline{u},\underline{y},\underline{m},t) &= \underline{0} \\ \underline{Q}(\underline{x},\underline{u},\underline{y},\underline{m},t_f) &= \underline{0} \end{aligned} \tag{4}$$

Defining the multipliers $\underline{\lambda}(t)$ to handle the differential constraints, $\underline{\rho}(t)$ to handle the equality constraints, and a constant $\underline{\mu}$ to handle the terminal time constraints, this problem may be written

$$\min_{\underline{x},\underline{u},\underline{y},\underline{m},\underline{\lambda},\underline{\rho},\underline{\mu}} J = \int_{t_o}^{t_f} (\underline{\lambda}^T\dot{\underline{x}}+H)\,dt + G \tag{5}$$

where the Hamiltonian is defined

$$H \triangleq f - \underline{\lambda}^T\underline{g} + \underline{\rho}^T\underline{q} \tag{6}$$

and

$$G \triangleq F(\underline{x}(t_f)) + \underline{\mu}^T\underline{Q} \tag{7}$$

If standard optimal control theory is applied, the following conditions of optimality are obtained:

$$\begin{aligned} \dot{\underline{x}} - \underline{g} &= \underline{0} \qquad \underline{x}(t_o) = \underline{x}_o \\ \underline{q} &= \underline{0} \\ \underline{Q}(t_f) &= \underline{0} \\ \dot{\underline{\lambda}} - \underline{H}_x &= \underline{0} \\ \underline{H}_u &= \underline{0} = \underline{H}_y = \underline{H}_m \\ (\underline{\lambda}+\underline{G}_x)_{t_f} &= \underline{0} \end{aligned} \tag{8}$$

subject to the inequality constraints at each interval. The object is to find controls and states satisfying these equations to a specified degree of accuracy. Defining a norm function

$$N(\underline{\xi}) = \underline{\xi}^T\underline{\xi} = \|\underline{\xi}\|^2 \tag{9}$$

we may specify the optimal control problem in terms of the minimization of the cumulative errors

$$J = J_1 + J_2 \tag{10}$$

where

$$J_1 = \int_{t_o}^{t_f} \{N(\underline{\dot{x}}-\underline{g}) + N(\underline{q})\}dt + N(\underline{Q}(t_f)) \tag{11}$$

$$J_2 = \int_{t_o}^{t_f} \{N(\underline{\dot{\lambda}}-\underline{H}_x) + N(\underline{H}_x) + N(\underline{H}_y) + N(\underline{H}_m)\}dt + N(\underline{\lambda}+\underline{G}_x)_{t_f} \tag{12}$$

and stop the iterations when J is smaller than some pre-specified accuracy. The normal quasilinearization algorithm corresponds to updating the current $\underline{x},\underline{u},\underline{y},\underline{m},\underline{\lambda},\underline{\rho},\underline{\mu}$ by increments derived from the solution of the first order expansions of the optimality conditions (8). A new expansion point is then obtained which reduces J, and the procedure repeated until convergence is achieved.

The inequality constraints may in general be handled by penalty methods with very little modification of the quasilinearization method. In this present work, however, attention is restricted to what is the most important case in engineering problems, simple upper and lower physical bounds on the variables. In this case, a gradient projection method due to Rosen[8] may be used. This basically checks the variables at each iteration to see if the limits are violated. If a variable exceeds the limit, it is set equal to that limit, and the next step in the minimization proceeds along the projection of the gradient along that limit.

The object of this paper, then, is to obtain solutions to very general subsystem problems and to coordinate their solutions into a coordination algorithm for the hierarchical control of the system.

3. Quasilinearization algorithm

The standard method of solving the problem outlined in the previous section is to consider the linearized set of optimality conditions defined by

$$
\begin{aligned}
(\dot{\underline{x}}-\underline{g})^k + \delta(\dot{\underline{x}}-\underline{g}) &= \underline{0} \\
\delta\underline{x}_o &= \underline{0} \\
\underline{q}^k + \delta\underline{q} &= \underline{0} \\
\underline{Q}^k + \delta\underline{Q}(t_f) &= \underline{0} \\
(\dot{\underline{\lambda}}-\underline{H}_x)^k + \delta(\dot{\underline{\lambda}}-\underline{H}_x) &= \underline{0} \\
\underline{H}_u^k + \delta\underline{H}_u &= \underline{0} \\
\underline{H}_y^k + \delta\underline{H}_y &= \underline{0} \\
\underline{H}_m^k + \delta\underline{H}_m &= \underline{0} \\
\delta(\underline{\lambda}+\underline{G}_x)_{t_f} + (\underline{\lambda}+\underline{G}_x)_f^k &= \underline{0}
\end{aligned} \tag{13}
$$

where the linearization is performed about some nominal operating point at the k^{th} iteration, denoted by the superscript k . These may be expanded to yield the following set of differential equations:

$$
\begin{aligned}
&\Delta\dot{\underline{x}} - \underline{g}_x^T\Delta\underline{x} - \underline{g}_u^T\Delta\underline{u} - \underline{g}_y^T\Delta\underline{y} - \underline{g}_m^T\Delta\underline{m} + (\dot{\underline{x}}-\underline{g})^k = \underline{0} \\
&\Delta\dot{\underline{\lambda}} - \underline{H}_{xx}^T\Delta\underline{x} - \underline{H}_{xu}^T\Delta\underline{u} - \underline{H}_{xy}^T\Delta\underline{y} - \underline{H}_{xm}^T\Delta\underline{m} - \underline{H}_{x\rho}^T\Delta\underline{\rho} + (\dot{\underline{\lambda}}-\underline{H}_x)^k - \underline{H}_{x\lambda}^T\Delta\underline{\lambda} = 0
\end{aligned} \tag{14}
$$

subject to the algebraic constraints

$$
\begin{aligned}
&\underline{q}_x^T\Delta\underline{x} + \underline{q}_u^T\Delta\underline{u} + \underline{q}_y^T\Delta\underline{y} + \underline{q}_m^T\Delta\underline{m} + \underline{q}^k = \underline{0} \\
&\underline{H}_{ux}^T\Delta\underline{x} + \underline{H}_{uu}^T\Delta\underline{u} + \underline{H}_{uy}^T\Delta\underline{y} + \underline{H}_{um}^T\Delta\underline{m} + \underline{H}_{u\lambda}^T\Delta\underline{\lambda} + \underline{H}_{u\rho}^T\Delta\underline{\rho} + \underline{H}_u^k = \underline{0} \\
&\underline{H}_{yx}^T\Delta\underline{x} + \underline{H}_{yu}^T\Delta\underline{u} + \underline{H}_{yy}^T\Delta\underline{y} + \underline{H}_{ym}^T\Delta\underline{m} + \underline{H}_{y\lambda}^T\Delta\underline{\lambda} + \underline{H}_{y\rho}^T\Delta\underline{\rho} + \underline{H}_y^k = \underline{0} \\
&\underline{H}_{mx}^T\Delta\underline{x} + \underline{H}_{mu}^T\Delta\underline{u} + \underline{H}_{my}^T\Delta\underline{y} + \underline{H}_{mm}^T\Delta\underline{m} + \underline{H}_{m\lambda}^T\Delta\underline{\lambda} + \underline{H}_{m\rho}^T\Delta\underline{\rho} + \underline{H}_m^k = \underline{0}
\end{aligned} \tag{15}
$$

and the initial and terminal conditions

$$
\begin{aligned}
&(\Delta\underline{x})_{t_o} = \underline{0} \\
&(\underline{Q}_x^T\Delta\underline{x})_{t_f} + (\underline{Q}_u^T\Delta\underline{u})_{t_f} + (\underline{Q}_y^T\Delta\underline{y})_{t_f} + (\underline{Q}_m^T\Delta\underline{m})_{t_f} + \underline{Q}(t_f) = \underline{0} \\
&(\Delta\underline{\lambda})_{t_f} + \{\underline{G}_{xx}^T\Delta\underline{x} + \underline{G}_{xu}^T\Delta\underline{u} + \underline{G}_{xy}^T\Delta\underline{y} + \underline{G}_{xm}^T\Delta\underline{m} + \underline{G}_{x\mu}^T\Delta\underline{\mu}\}_{t_f} + (\underline{\lambda}+\underline{G}_x)_{t_f} = \underline{0}
\end{aligned} \tag{16}
$$

The solutions $\Delta\underline{x},\Delta\underline{\lambda},\Delta\underline{u},\Delta\underline{m},\Delta\underline{y},\Delta\underline{\mu},\Delta\underline{\rho}$ may then be added to the current estimate of the solution to yield an estimate at the k^{th} iteration of

$$\begin{aligned}
\underline{x}^{k+1} &= \underline{x}^{k} + \Delta\underline{x} \\
\underline{\lambda}^{k+1} &= \underline{\lambda}^{k} + \Delta\underline{\lambda} \\
\underline{u}^{k+1} &= \underline{u}^{k} + \Delta\underline{u} \\
\underline{m}^{k+1} &= \underline{m}^{k} + \Delta\underline{m} \\
\underline{y}^{k+1} &= \underline{y}^{k} + \Delta\underline{y} \\
\underline{\mu}^{k+1} &= \underline{\mu}^{k} + \Delta\underline{\mu} \\
\underline{\rho}^{k+1} &= \underline{\rho}^{k} + \Delta\underline{\rho}
\end{aligned} \tag{17}$$

This leads to a change in the performance index of

$$\delta J_1 + \delta J_2 \tag{18}$$

where

$$\delta J_1 = 2\int_{t_o}^{t_f} \{(\dot{\underline{x}}-\underline{g})^T \delta(\dot{\underline{x}}-\underline{g}) + \underline{q}^T \delta\underline{q}\}dt + 2\underline{Q}^T(t_f)\delta\underline{Q}(t_f) \tag{19}$$

$$= -2\,J_1 \tag{20}$$

$$\delta J_2 = 2\int_{t_o}^{t_f} \{(\dot{\underline{\lambda}}-\underline{H}_x)^T \delta(\dot{\underline{\lambda}}-\underline{H}_x) + \underline{H}_u^T \delta\underline{H}_u + \underline{H}_y^T \delta\underline{H}_y + \underline{H}_m^T \delta\underline{H}_m\}dt + 2(\underline{\lambda}+\underline{G}_x)^T_{t_f}\,\delta(\underline{\lambda}+\underline{G}_x)_{t_f} \tag{21}$$

$$= -2\,J_2 \tag{22}$$

substituting in for the incremental changes of equation (13) and using the definitions of J_1 and J_2 . Thus

$$\delta J = -2J \tag{23}$$

showing a reduction in the value of J at each iteration.

The modified[6] quasilinearization method uses the following linearized version of the optimality conditions for $0 \leq \alpha \leq 1$

$$\begin{aligned}
(\dot{\underline{x}}-\underline{g})^k + \alpha\delta(\dot{\underline{x}}-\underline{g}) &= \underline{0} \\
\alpha\delta\underline{x}_o &= \underline{0} \\
\underline{q}^k + \alpha\delta\underline{q} &= \underline{0} \\
\underline{Q}^k + \alpha\delta\underline{Q}(t_f) &= \underline{0} \\
(\dot{\underline{\lambda}}-\underline{H}_x)^k + \alpha\delta(\dot{\underline{\lambda}}-\underline{H}_x) &= \underline{0} \\
\underline{H}_u^k + \alpha\delta\underline{H}_u &= \underline{0} \\
\underline{H}_y^k + \alpha\delta\underline{H}_y &= \underline{0} \\
\underline{H}_m^k + \alpha\delta\underline{H}_m &= \underline{0} \\
(\underline{\lambda}+\underline{G}_x)_f^k + \alpha\delta(\underline{\lambda}+\underline{G}_x)_f^k &= \underline{0}
\end{aligned} \tag{24}$$

which, if expanded, leads to a reduction in the performance index

$$\delta J = -2\alpha J \tag{25}$$

and hence, by suitable choice of α at each iteration, the performance index may be rapidly extremized. This optimum choice of α may be obtained by a one-dimensional search on $J(\alpha)$ at each iteration.

4. Soltuon of two-point boundary value problem

The previous section has formulated the quasilinearization approach for a general control problem. Here, the solution algorithm for the linearized two-point boundary-value problem is described.

Consider the problem of solving the differential equations (14) subject to the algebraic constraints (15) and the initial and terminal conditions (16). Clearly it is possible to solved equations (15) for $\Delta\underline{u},\Delta\underline{y},\Delta\underline{m},\Delta\underline{\rho}$ all as functions of $\Delta\underline{x}$ and $\Delta\underline{\lambda}$, since then there are 4 equations in 4 unknowns. Substituting these values into equations (14) and defining

$$\underline{z}^{k+1} = \begin{bmatrix} \underline{x}^{k+1} \\ \underline{\lambda}^{k+1} \end{bmatrix} = \begin{bmatrix} \underline{x}^{k} + \Delta\underline{x} \\ \underline{\lambda}^{k} + \Delta\underline{\lambda} \end{bmatrix} \tag{26}$$

the problem may be restated as the solution of the linearized equation

$$\frac{d}{dt}\underline{z}^{k+1} = \underline{F}(\underline{z}^{k}) + \underline{J}(\underline{z}^{k})(\underline{z}^{k+1}-\underline{z}^{k}) \tag{27}$$

subject to the boundary conditions

$$\begin{aligned} z_i(t_o) &= z_{io} \qquad i = 1,2,\ldots,n \\ z_i(t_f) &= z_{if} \qquad i = n+1,\ldots,2n \end{aligned} \tag{28}$$

where z_i is the i^{th} component of the $2n\times1$ vector $\underline{z}$, and $\underline{J}$ is the system Jacobian with ij^{th} element

$$J_{ij} = \frac{\partial F_i}{\partial z_j} \tag{29}$$

It is convenient at this stage to write the boundary conditions as

$$\begin{aligned} z_1(t_o)a_{k_1} + z_2(t_o)a_{k_2} + \ldots z_{2n}(t_o)a_{k_{2n}} &= b_k \\ z_1(t_f)a_{\ell_1} + z_2(t_f)a_{\ell_2} + \ldots z_{2n}(t_f)a_{\ell_{2n}} &= b_\ell \end{aligned} \tag{30}$$

where there are k initial conditions and ℓ final conditions on $z_i(t)$.

Writing equation (27) as

$$\frac{d}{dt}\underline{z}^{k+1} = \underline{J}(\underline{z}^k)\underline{z}^{k+1} + \underline{F}(\underline{z}^k) - \underline{J}(\underline{z}^k)\underline{z}^k \tag{31}$$

it is clear that the general solution consists of two parts

$$\underline{z}^{k+1}(t) = \underline{\Phi}^{k+1}(t,t_o)\underline{z}^{k+1}(t_o) + \underline{P}(t) \tag{32}$$

The first part or transient response is determined by solving

$$\frac{d}{dt}\underline{\Phi}^{k+1}(t,t_o) = \underline{J}(\underline{z}^k)\underline{\Phi}^{k+1}(t,t_o) \tag{33}$$

subject to

$$\underline{\Phi}^{k+1}(t_o,t_o) = \underline{I} \tag{34}$$

where $\underline{I}$ is the $2n\times 2n$ identity matrix. The particular integral satisfies

$$\frac{d}{dt}\underline{P}^{k+1}(t) = \underline{J}(\underline{z}^k)\underline{P}^{k+1}(t) + \underline{F}(\underline{z}^k) - \underline{F}(\underline{z}^k)\underline{z}^k \tag{35}$$

and the general solution is

$$\underline{z}^{k+1}(t) = \underline{\Phi}^{k+1}(t,t_o)\underline{C}^{k+1} + \underline{P}^{k+1}(t)$$

where $\underline{C}^{k+1}$ is the $2n\times 1$ vector of constants of integration determined by the equations (30) which may be rewritten in the form

$$\begin{aligned} <\underline{C}^{k+1},\underline{a}_i> &= b_i \qquad i = 1,2,\ldots n \\ <\underline{\Phi}^{k+1}(t_f,t_o)\underline{C}^{k+1}+\underline{P}^{k+1}(t_f),\underline{a}_j> &= b_j \qquad j = n+1,\ldots 2n \end{aligned} \tag{36}$$

where

$$\underline{a}_i = [a_{i_1},a_{i_2},\ldots,a_{i_{2n}}]^T \tag{37}$$

The iterations continue until

$$\| \underline{c}^{k+1} - \underline{c}^{k} \| \leqslant \varepsilon \quad (38)$$

where ε is an arbitrarily small quantity. The basic steps of the solution algorithm are shown on the flowchart of Fig.1.

The following section provides an illustration of the practical use of this technique for solving non-linear sub-system problems in a hierarchical structure arising from the spatial problem decomposition using Lagrangian duality.

5. Solution of a non-linear water supply problem

Consider the problem of obtaining the minimum pumping cost control for a water supply system[9] described by the following differential equations:

$$\frac{dx_1}{dt} = 0.075\times u_1 - 4.66\times 10^{-4}\times S_1(x_1-x_2) - 2.367\times 10^{-3} \quad (39)$$

$$\frac{dx_2}{dt} = 9.527\times 10^{-4}\times S_1(x_1-x_2) + 0.1533\times u_2 - 2.485\times 10^{-3}\times S_2(x_2-x_5) - 3.629\times 10^{-3} \quad (40)$$

$$\frac{dx_3}{dt} = 0.1\times u_3 - 0.0239 \quad (41)$$

$$\frac{dx_4}{dt} = 0.25\times u_4 - 3.726\times 10^{-3}\times S_3(x_4-239)^{0.54} - 2.95\times 10^{-3} \quad (42)$$

$$\frac{dx_5}{dt} = 0.05\times u_5 + 8.1\times 10^{-4}\times S_2(x_2-x_5) - 8.97\times 10^{-3} \quad (43)$$

$$\frac{dx_6}{dt} = 0.1712\times u_6 + 1.25\times 10^{-3}\times S_4(375-x_6) + 9.922\times 10^{-4}\times S_5(540-x_6) - 1.35\times 10^{-3} \quad (44)$$

subject to $\underline{x}(0) = \underline{x}_o$, where

$$J = \int_0^{24} \{\frac{1}{2} \| \underline{x}-\underline{x}^d \|^2_{\underline{Q}} + \underline{Q}_L\underline{x} + \frac{1}{2} \| \underline{u}-\underline{u}^d \|^2_{\underline{R}} + \underline{R}_L\underline{u}\}dt \quad (45)$$

is a second-order expansion of operating costs. In this problem $\underline{x}$ is the vector of states or reservoir levels on the system, $\underline{u}$ is the vector of pump controls, and $\underline{x}^d, \underline{u}^d$ are their desired values respectively. All vectors are of dimension 6 , and $\underline{Q}, \underline{Q}_L, \underline{R}, \underline{R}_L$ are the appropriate 6×6 weighting matrices.

This regulator control problem arises from an actual control scheme developed on a computer-controlled water supply network, described in detail in reference 9.

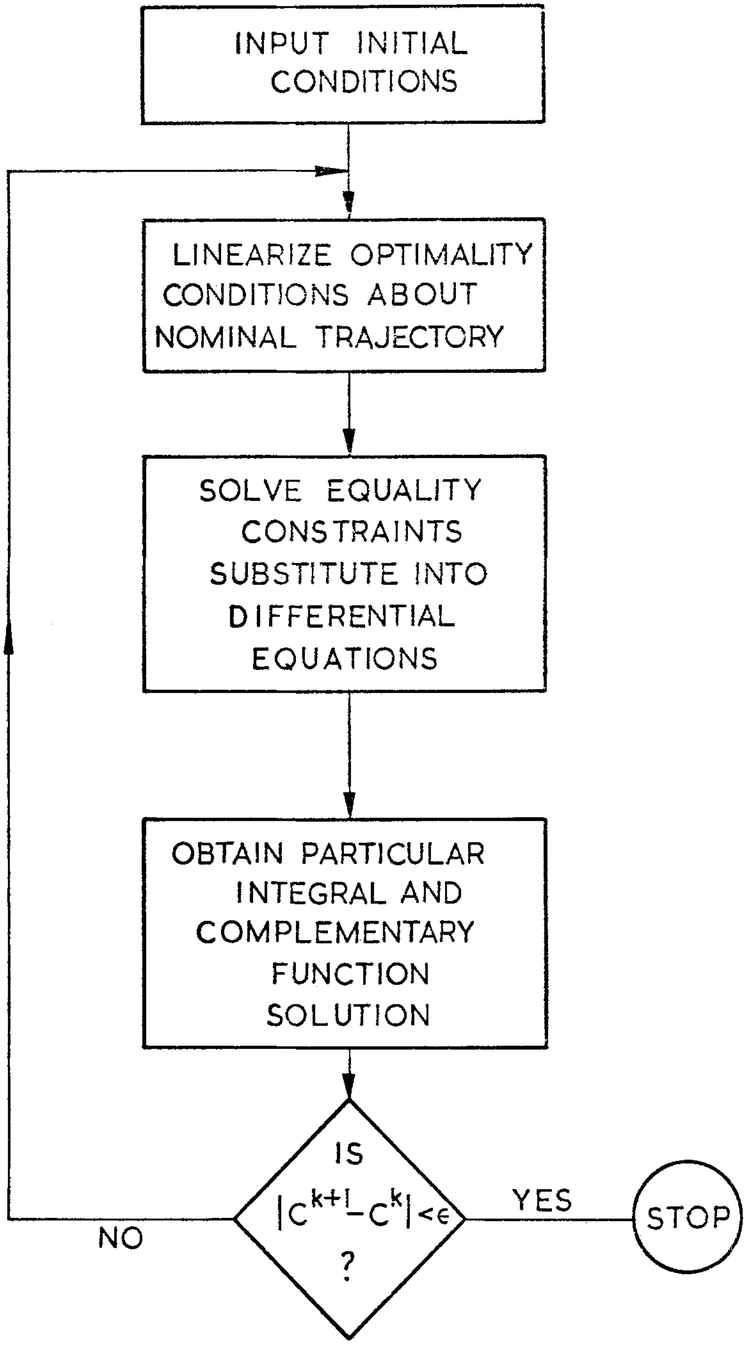

FIG.1 QUASILINEARISATION ALGORITHM FOR SOLVING SUB-SYSTEM PROBLEMS

The non-linearities $S_1 \dots S_6$ appearing in this formulation are of the general form

$$S_i(\zeta) = \mathrm{Sgn}(\zeta)|\zeta|^{0.54} \qquad i = 1,2,\dots 5 \qquad (46)$$

which is in fact the steady state flow along a pipe of unit resistance, the head drop along the pipe being ζ . A graph of this is shown in Fig.2.

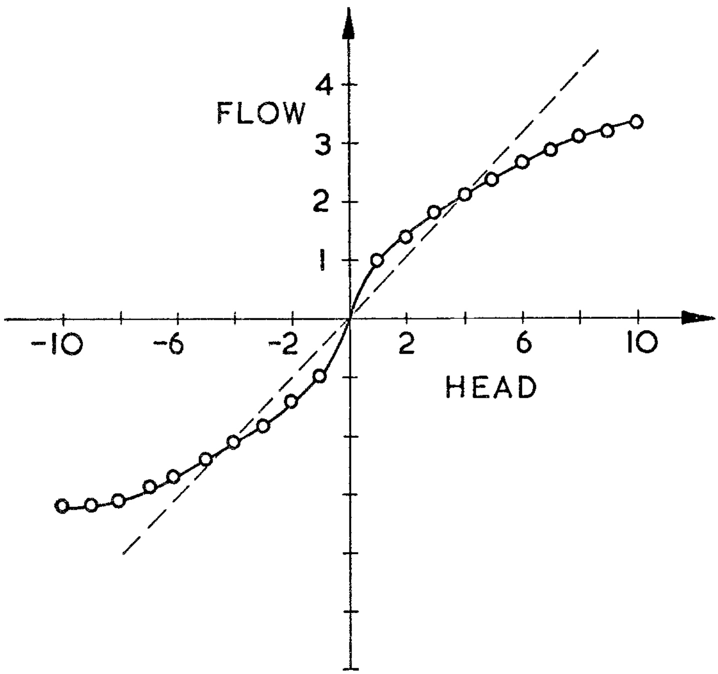

FIG.2 STEADY-STATE HEAD-FLOW RELATIONSHIP FOR A PIPE

This problem may be decomposed by introducing the coupling variables

$$\begin{aligned} \xi_{11} &= x_2 \\ \xi_{21} &= x_1 \\ \xi_{22} &= x_5 \\ \xi_{51} &= x_2 \end{aligned} \qquad (47)$$

where ξ_{ij} is the j^{th} coupling variable in the i^{th} sub-system. In this way the following 6 sub-problems are defined:

Sub-problem 1:

$$\min_{x_1,u_1,\xi_{11}} \int_0^{24} \{\frac{1}{2} u_1^2 + \frac{1}{2}(x_1-765)^2 + \beta_{11}(\xi_{11}-x_2)\}dt + F_1(x_1(t_f))$$

S.t.
$$\dot{x}_1 = 0.075\times u_1 - 4.66\times 10^{-4}\times S_1(x_1-\xi_{11}) - 2.367\times 10^{-3}$$
$$x_1(0) = 765$$
$$u_1(t) \leqslant 1$$

Sub-problem 2:

$$\min_{x_2,u_2,\xi_{21},\xi_{22}} \int_0^{24} \{\frac{1}{2}(x_2-535)^2 + \frac{1}{2} u_2^2 + \beta_{21}(\xi_{21}-x_1) + \beta_{22}(\xi_{22}-x_5)\}dt + F_2(x_2(t_f))$$

S.t.
$$\dot{x}_2 = 9.527\times 10^{-4}\times S_1(\xi_{21}-x_2) + 0.1533\times u_2 - 2.48\times 10^{-3}\times S_2(x_2-\xi_{22}) - 3.629\times 10^{-3}$$
$$x_2(0) = 535$$
$$u_2(t) \leqslant 1$$

Sub-problem 3:

$$\min_{x_3,u_3} \int_0^{24} \{\frac{1}{2}(x_3-489)^2 + \frac{1}{2} u_3^2\}dt + F_3(x_3(t_f))$$

S.t.
$$\dot{x}_3 = 0.1\times u_3 - 0.0239$$
$$x_3(0) = 489$$
$$u_3(t) \leqslant 1$$

Sub-problem 4:

$$\min_{x_4,u_4} \int_0^{24} \{\frac{1}{2}(x_4-307)^2 + \frac{1}{2} u_4^2\}dt + F_4(t_f))$$

S.t.
$$\dot{x}_4 = 0.25\times u_4 - 3.726\times 10^{-3}\times S_3(x_4-239) - 2.95\times 10^{-3}$$
$$x_4(0) = 307$$
$$u_4(t) \leqslant 1$$

Sub-problem 5:

$$\min_{x_5,u_5,\xi_{51}} \int_0^{24} \{\frac{1}{2}(x_5-393)^2 + \frac{1}{2} u_5^2 + \beta_{51}(\xi_{51}-x_2)\}dt + F_5(x_5(t_f))$$

S.t.
$$\dot{x}_5 = 0.05\times u_5 + 8.1\times 10^{-4}\times S_2(\xi_{51}-x_5) - 8.97\times 10^{-3}$$
$$x_5(0) = 393$$
$$u_5(t) \leqslant 1$$

Sub-problem 6:

$$\underset{x_6,u_6}{\text{Min}} \int_0^{24} \{\tfrac{1}{2}(x_6-245)^2 + \tfrac{1}{2}u_6^2\}dt + F_6(x_6(t_f))$$

$$\text{S.t.} \quad \dot{x}_6 = 0.1712\times u_6 + 1.25\times10^{-3}\times S_4(375-x_6) + 9.922\times10^{-4}\times S_5(540-x_6) - 1.35\times10^{-3}$$

$$x_6(0) = 245$$

$$u_6(t) \leqslant 1$$

Using a goal coordination algorithm, each of these sub-problems is solved for a given value of the coupling multiplier $\underline{\beta}$. The solution $\underline{x}_i^*$, $\underline{u}_i^*$ and $\underline{\xi}_i^*$ of the sub-system problems are then used to update the multiplier according to

$$\underline{\beta}^{k+1} = \underline{\beta}^k + \alpha\underline{d}^k \tag{48}$$

where

$$\underline{d}^k = \nabla_{\underline{\beta}}\psi(\underline{\beta}) \tag{49}$$

for a linear search or

$$\underline{d}^k = [\nabla_{\underline{\beta}}\psi(\underline{\beta})]^k + \frac{\|\nabla_{\underline{\beta}}\psi(\underline{\beta})\|_k^2}{\|\nabla_{\underline{\beta}}\psi(\underline{\beta})\|_{k-1}^2}\underline{d}^{k-1} \tag{50}$$

for a conjugate gradient minimization of the dual function $\psi(\underline{\beta})$. In this case, the gradient is simply the error in the coordinating relations, and when this tends to zero, the solution has been obtained. In this case the dual function is

$$\psi(\underline{\beta}) = \underset{\underline{x}_i,\underline{u}_i,\underline{\xi}_i}{\text{Min}} \{\sum_{i=1}^{6} [F_i + \int_{t_o}^{t_f} (f_i + \langle\underline{\beta},\underline{G}_i\rangle)dt]\} \tag{51}$$

and from this

$$\nabla_{\underline{\beta}}\psi(\underline{\beta}) = \sum_{i=1}^{6} \underline{G}_i \tag{52}$$

for the minimizing values of $\underline{x}_i,\underline{u}_i,\underline{\xi}_i$.

Minimizing the Hamiltonian H_i for each sub-system, as in Section 2, leads to the optimality conditions

$$\dot{\underline{x}}_i = \frac{\partial H_i}{\partial\underline{\lambda}_i}$$

$$\dot{\underline{\lambda}}_i = -\frac{\partial H_i}{\partial\underline{x}_i} \tag{53}$$

$$\frac{\partial H_i}{\partial \underline{u}_i} = \underline{0}$$
$$\frac{\partial H_i}{\partial \underline{\xi}_i} = \underline{0} \qquad (54)$$

The values of $\underline{u}_i$ and $\underline{\xi}_i$ in terms of $\underline{x}_i$ and $\underline{\lambda}_i$ may be solved for from equations (54) and the problem reduces to the solution of the two-point boundary-value problem defined by equations (53), subject to the initial conditions

$$\underline{x}_i(0) = \underline{x}_{io} \qquad \underline{\lambda}_i(t_f) = \underline{\lambda}_{if}$$

For all of the sub-problems, 24 integration steps are used, and the integration of the differential equations performed using a 4^{th} order Runge-Kutta procedure.

6. Results

The above problem has been solved on an IBM 370/165 computer, and a typical system response to a step disturbance for the first sub-problem is shown in Fig.3. The described quasilinearization algorithm has been used to obtain sub-system solutions. The convergence requirements for each sub-system are summarized in Table 1, together with the initial values of the multipliers. The global problem converges in about 30 iterations using a simple linear search.

The optimum controls are calculated to minimize the control effort and the deviations of the state from desired levels when a step demand is placed on the system. As shown in Fig.3, a step demand for water is met by reducing the pump flow control as much as possible and meeting the difference by reducing the reservoir level. For example, the final control for sub-system 1 is

$$u_1^*(24) = 0.075 \times u_1(24) = 2.367 \times 10^{-3} \text{ mgd} \qquad (55)$$

and this is the minimum allowable pumped quantity of water.

In general, this zone will be subjected to a time varying deterministic disturbance which is obtained by forecasting the water demand from past data. Fig.3 merely shows the response to the steady state or average of this demand.

An important practical consideration in the proposed on-line implementation of a scheme designed to handle non-linear models is whether the increased computational burden is justifiable in terms of cost savings.

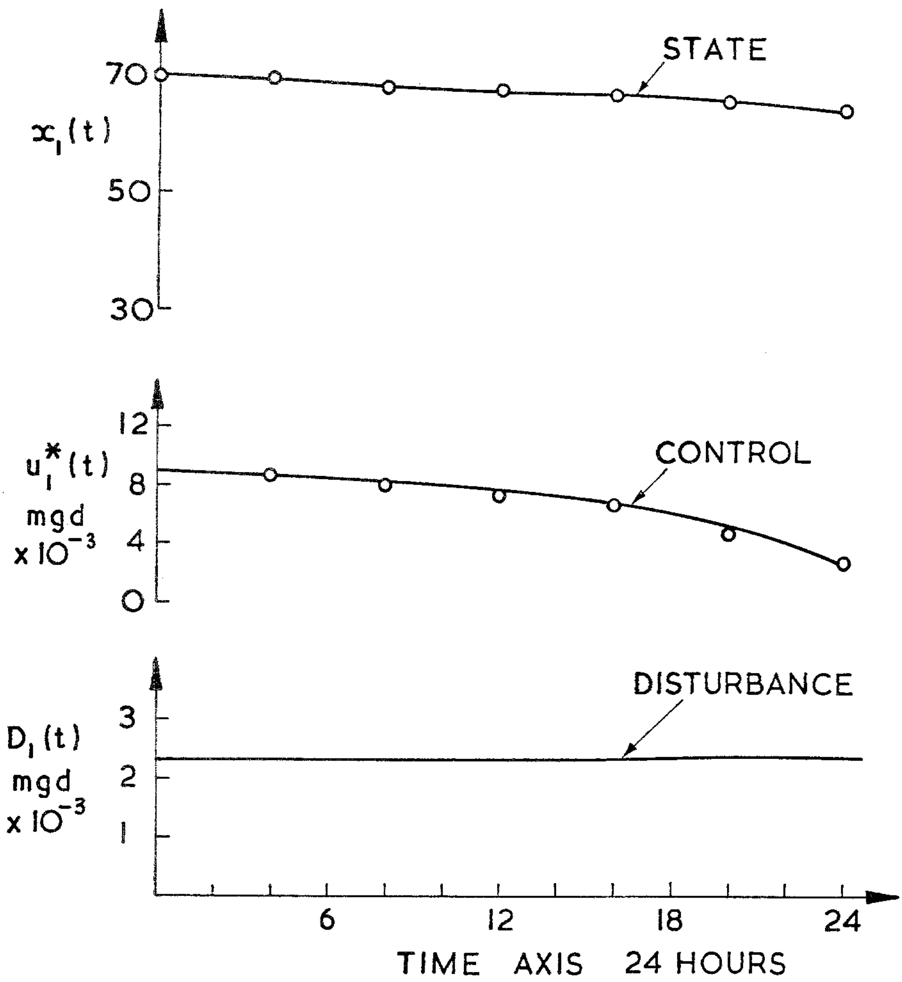

FIG.3 OPTIMAL CONTROL AND STATE FOR SUB-PROBLEM 1

Table 1

Sub-problem i	No. of iterations* for convergence	Initial multipliers $\lambda_i(t)$ $0 \leqslant t \leqslant 24$	$\lambda_i(24)$
1	3	0.4	0.4
2	2	0.15	0.15
3	2	2.39	2.39
4	12	0.047	0.047
5	3	3.58	3.58
6	20	0.0457	0.0457

*iterations are terminated when

$$\|c^{k+1}-c^k\| \leqslant 10^{-6}$$

To examine this, consider the solution to the first sub-system problem, using firstly a linear model and secondly a non-linear model. In the first case, the model is linearized about an operating point and the quadratic costs calculated. The results are tabulated in Table 2. Then, for a full non-linear model, solved by quasilinearization, the trajectories are obtained and quadratic costs calculated, the results being shown in Table 3.

Table 2

t	$\Delta x_1 = 765-x_1$	$(\Delta x_1)^2$	$\lambda_1(t)$	$\lambda_1^2(t)$
4	9.766×10^{-3}	95.374×10^{-6}	1.489	2.21855
8	21.729×10^{-3}	472.149×10^{-6}	1.429	2.04209
12	36.621×10^{-3}	134.109×10^{-5}	1.312	1.746283
16	53.711×10^{-3}	288.487×10^{-5}	1.146	1.31265
20	75.928×10^{-3}	576.506×10^{-5}	0.886	0.78503
24	104.98×10^{-3}	110.208×10^{-4}	0.513	0.26273

$$\sum_t (\Delta x_1)^2 = 21579.352\times10^{-6}$$

$$\sum_t \lambda_1^2 = 8.3677339$$

Table 3

t	$\Delta x_1 = 765-x_1$	$(\Delta x_1)^2$	$\lambda_1(t)$	$\lambda_1^2(t)$
4	9.521×10^{-3}	90.649×10^{-6}	1.491	2.22243
8	21.729×10^{-3}	472.149×10^{-6}	1.430	2.04622
12	36.377×10^{-3}	132.328×10^{-5}	1.323	1.75080
16	53.467×10^{-3}	285.872×10^{-5}	1.140	1.29969
20	75.195×10^{-3}	565.43×10^{-5}	0.880	0.776108
24	104.492×10^{-3}	109.185×10^{-4}	0.512	0.26245

$$\sum_t (\Delta x_1)^2 = 21317.67\times10^{-6}$$

$$\sum_t (\lambda_1)^2 = 8.357703$$

The quadratic costs in the first case are

$$\begin{aligned} J_1 &= \sum_t (\Delta x_1)^2 + \sum_t (u_1)^2 \\ &= \sum_t (\Delta x_1)^2 + (0.075)^2 \sum_t (\lambda_1)^2 \\ &= 686.546\times10^{-4} \end{aligned}$$

and in the second case

$$J_2 = 683.297\times10^{-4}$$

which is a percentage reduction of about 0.5% . The computational burden in terms of time requirements for handling the non-linearity is increased by a factor of 3 . Clearly the increase depends on the number of iterations required in total to solve all the sub-system problems, and these can be reduced to a minimum using modifications of the quasilinearization algorithm. Experience has shown that the number of iterations to convergence can be halved using the modified algorithm and optimizing with respect to the choice of initial multipliers over the time interval.[6] Thus, for a 50% increase in computer time requirements, pumping costs can be reduced by ½% .

This additional saving of ½% in costs may be well worthwhile if the dynamics of the system under consideration are slow enough to allow the computer to calculate the optimal controls within one sampling interval, as is the case with the water supply problem. If, however, the additional time requirements of the quasilinearization algorithm mean that a more powerful computer must be installed to control the system, then the savings will have to be compared with the additional computer costs to determine if it is worthwhile increasing the computer capacity to save the additional ½% operating costs.

Conclusions

The method of quasilinearization has been used to solve non-linear sub-system problems arising from a problem decomposition using Lagrangian duality. The method has been applied to a practical water supply problem, and results indicate that it is worthwhile handling non-linear models for on-line implementation in this application.

Acknowledgements

The authors are grateful to R.H. Burch and K.C. Marlow, of East Worcestershire Waterworks Company, for financial support, and to A.R. Farmer and M.S. Jennions of Kent Automation Systems, for technical cooperation. They are also grateful to M. Singh and J. Galy of Laboratoire d'Automatique et d'Analyse des Systemes du C.N.R.S., for a computer program on which these calculations are based, and to

P.D. Rice, of this department, for data processing. Thanks are also due to the Science Research Council and Peterhouse, Cambridge, for financial support.

References

1. R. Bellman: "Dynamic Programming", Princeton University Press, Princeton, New Jersey, 1957 (book)

2. R. Bellman and S. Dreyfus: "Applied Dynamic Programming", Princeton University Press, Princeton, New Jersey, 1962 (book).

3. R.A. Howard: "Dynamic Programming and Markov Processes", John Wiley and Sons Inc., New York, 1960 (book)

4. R.E. Larson: "State Increment Dynamic Programming", John Wiley and Sons Inc., New York, 1960 (book)

5. R. Bellman and R.E. Kalaba: "Quasilinearization and Nonlinear Boundary-Value Problems", American Elsevier Publishing Company, Inc., New York, 1965 (book)

6. A. Miele, A. Mangiavacchi and A.K. Aggarwal: "Modified Quasilinearization Algorithm for Optimal Control Problems with Nondifferential Constraints", Journal of Optimization Theory and Applications, Vol.14, No.5, 1974

7. F. Fallside and P.F. Perry: "Decentralized Optimum Control Methods for Water Distribution System's Optimization", Cambridge University Engineering Dept., TR 31 (elec) 1974

8. J.B. Rosen: "The Gradient Projection Method for Nonlinear Programming - Pts 1 & 2", Journal Soc.Ind.Appl.Math., 1960, 8, pp.181-217

9. F. Fallside and P.F. Perry: "Hierarchical Optimization of a Water-Supply Network", Proc.IEE, Vol.122, No.2, Feb. 1975, pp.202-208