OPTIMIZATION TECHNIQUES FOR MINIMIZATION OF COMBINED SEWER OVERFLOW

prepared by

John W. Labadie

Metropolitan Water Intelligence Systems

Technical Report No. 11

June 1973

Water Resource Systems Program
Department of Civil Engineering
Colorado State University
Fort Collins, Colorado

FOREWORD

This is one of a series of Technical Reports prepared under a grant by the Office of Water Resources Research which supports a project at Colorado State University entitled "Metropolitan Water Intelligence Systems." The objective of the project is to develop criteria and information for the development of metropolitan water intelligence systems (MWIS). The MWIS is a specialized form of the management information and control system concept which is becoming popular as a tool in industrial applications.

The project consists of three phases, each lasting about one year. This report was prepared during Phase II. Basic objectives for Phase I were to:

1. Investigate and describe modern automation and control systems for the operation of urban water facilities with emphasis on combined sewer systems.
2. Develop criteria for managers, planners, and designers to use in the consideration and development of centralized automation and control systems for the operation of combined sewer systems.
3. Study the feasibility, both technical and social, of automation and control systems for urban water facilities with emphasis on combined sewer systems.

Basic objectives for Phase II are:

1. Formulate a design strategy for the automation and control of combined sewer systems.
2. Develop a model of a real-time automation and control system (RTACS model).
3. Describe the requirements for computer and control equipment for automation and control systems.
4. Describe nontechnical problems associated with the implementation of automation and control systems.

This report concentrates on methods of developing control logic for automated operation of ambient and/or auxiliary storage capabilities within combined sewer systems, with the objective of minimizing overflows to receiving waters. The enormous number of control opportunities requires that the control problem be formulated as an optimization problem. The problem is defined as one of minimizing total weighted overflows, subject to an assumed hydraulic model describing flow and storage dynamics, as well as other physical constraints. The optimization problem tends to increase in complexity and degree of nonlinearity as less idealized flow models are utilized. This report concentrates on limited subbasin analysis, with the view that the large-scale problem is ultimately solved by a master control algorithm that ties the subbasins together in an iterative fashion.

Finite-dimensional optimization techniques appear to have greater potential for effective solution, over infinite-dimensional techniques (i.e., application of continuous-time optimal control theory). The primary reasons are (i) difficulty of obtaining solutions by the latter, (ii) operation of the real system in discrete time. Within the category of finite-dimensional optimization, indirect solution of the optimization problem through application of generalized duality theory has greater potential for finding global solutions than direct application of mathematical programming techniques. This is made possible through development of an *approximate-flow* technique that significantly reduces the total number of variables involved in the problem. Considerable off-line computational work is required to fully verify these assertions.

* * * * *

This report was supported by OWRR grant number 14-31-0001-3685, Title II, Project No. C-3105, from funds provided by the United States Department of Interior as authorized under the Water Resources Research Act of 1964, Public Law 88-379, as amended.

* * * * *

The following technical reports were prepared during Phase I of the CSU-OWRR project, Metropolitan Water Intelligence Systems. Copies may be obtained for $3.00 from the National Technical Information Service, U. S. Department of Commerce, Springfield, VA 22151. (When ordering, use the report title and the identifying number noted for each report.)

Technical Report No. 1 - "Existing Automation, Control and Intelligence Systems of Metropolitan Water Facilities" by H. G. Poertner. (PB 214266)

Technical Report No. 2 - "Computer and Control Equipment" by Ken Medearis. (PB 212569)

Technical Report No. 3 - "Control of Combined Sewer Overflows in Minneapolid-St. Paul" by L. S. Tucker. (PB 212903)

Technical Report No. 4 - "Task 3 - Investigation of the Evaluation of Automation and Control Schemes for Combined Sewer Systems" by J. J. Anderson, R. L. Callery, and D. J. Anderson. (PB 212573)

Technical Report No. 5 - "Social and Political Feasibility of Automated Urban Sewer Systmes" by D. W. Hill and L. S. Tucker. (PB 212574)

Technical Report No. 6 - "Urban Size and Its Relation to Need for Automation and Control" by Bruce Bradford and D. C. Taylor. (PB 212523)

Technical Report No. 7 - "Model of Real-Time Automation and Control Systems for Combined Sewers" by Warren Bell, C. B. Winn and George L. Smith. (PB 212575)

Technical Report No. 8 - "Guidelines for the Consideration of Automation and Control Systems" by L. S. Tucker and D. W. Hill. (PB 212576)

Technical Report No. 9 - "Research and Development Needs in Automation and Control of Urban Water Systems" by H. G. Poertner. (PB 212577)

Technical Report No. 10 - "Planning and Wastewater Management of a Combined Sewer System in San Francisco" by Neil S. Grigg, William R. Giessner, Robert T. Cockburn, Harold C. Coffee, Jr., Frank H. Moss, Jr., and Mark E. Noonan. (PB#-to be assigned)

Technical Report No. 11 - "Optimization Techniques for Minimization of Combined Sewer Overflow" by John W. Labadie. (PB#-to be assigned)

MATHEMATICAL NOTATION

In this report, for notational convenience, no attempt is made to distinguish between column and row vectors. It is presumed that the reader can distinguish this for himself.

$x \in E^n$	a *vector* $x = (x_1,\ldots,x_n)$, or an element of ($\in$) n-dimensional *Euclidean space* E^n. If $x_i \geq 0$ $(i = 1,\ldots,n)$, then $x \in (E^n)^+$.
$f(x)$	a *vector-valued function* $f(x) = (f_1(x),\ldots,f_m(x))$, also denoted as $f:E^n \to E^m$ or $f(\cdot) \in E^m$.
$\nabla_x f(x)$	the *gradient* vector of f, or $\nabla_x f(x) = (\frac{\partial f}{\partial x_1},\ldots,\frac{\partial f}{\partial x_n})$.
$x_i(k)$	component of a *matrix* of numbers, also denoted as $x(k) \in E^n$ $(k = 1,\ldots,m)$, $x_i(\cdot) \in E^n$ $(i = 1,\ldots,n)$, or $x(\cdot) \in E^{m+n}$.
$x \cdot y$	*scalar product* of two vectors $x,y \in E^n$, where $x \cdot y = \sum_{i=1}^{n} x_i y_i$.
$\{x \mid P(x)\}$	*set* S of elements $x \in E^n$ satisfying some given property P, where S is a subset of E^n, or $S \subset E^n$.
$[a,b]$	set defined over *closed interval* $a \leq x \leq b$, $x,a,b \in E^n$.
(a,b)	set defined over *open interval* $a < x < b$, $x,a,b \in E^n$.
convex set	set S is *convex* if and only if for every $x,y \in S \subset E^n$, $(\alpha x + (1-\alpha)y) \in S$, for all $\alpha \in [0,1]$.
int(S)	the *interior* of the set $S \subset E^n$, or the largest *open* set contained in S.

$X \times Y$	*cartesian product* of two sets $X \subset E^m, Y \subset E^n$; or $(X \times Y) \subset E^{m+n}$
$\dot{x}(t)$	vector of *derivatives* $(\frac{dx_1(t)}{dt}, \ldots, \frac{dx_n(t)}{dt})$.
x^*	*global solution* of optimization problem $\min\{f(x) \mid x \in S\}$, where $S \subset E^n$, or $f(x^*) \leq f(x)$, for all $x \in S$.
x^o	*local solution* of optimization problem $\min\{f(x) \mid x \in S\}$, $S \subset E^n$, or $f(x^o) \leq f(x)$, for all $x \in S \cap N$; where $N = (x^o - \varepsilon, x^o + \varepsilon)$, for some scalar $\varepsilon > 0$.
$\triangleq$	equal by definition
=>	forward implication (*implies*)
<=	reverse implication (is *implied by*)
<=>	equivalence, or if and only if [*iff*]

TABLE OF CONTENTS

	Page
FOREWORD	i
MATHEMATICAL NOTATION	ii
I. INTRODUCTION	1
A. THE CONTROL PROBLEM	1
B. OFF-LINE VS. ON-LINE OPTIMIZATION	2
C. SYSTEM DECOMPOSITION	4
D. OBJECTIVES	6
II. FINITE AND INFINITE-DIMENSIONAL OPTIMIZATION	8
A. A RESERVOIR CONTROL PROBLEM	8
1. Discrete Time Case	8
2. Continuous-Time Case	10
3. Discussion	11
B. NECESSARY CONDITIONS FOR DISCRETE-TIME OPTIMAL CONTROL	13
C. NECESSARY CONDITIONS FOR CONTINUOUS-TIME OPTIMAL CONTROL	15
1. From Discrete-Time to Continuous-Time	16
2. Necessary Conditions	18
3. Solution Difficulties	18
D. SUMMARY AND DISCUSSION	21
III. APPLICATIONS OF CONTINUOUS-TIME CONTROL THEORY	24
A. INTRODUCTION	24
B. AMBIENT STORAGE	25
1. Two Reservoirs in Series	25
2. A Three-Reservoir Problem	31
C. AUXILIARY STORAGE	35
D. DISCUSSION	37
IV. MATHEMATICAL PROGRAMMING APPROACHES	39
A. INTRODUCTION	39
B. AN EXAMPLE THREE-RESERVOIR PROBLEM	40
C. DIRECT METHODS	44
1. Linear Programming	44
2. Dynamic Programming	45
3. Nonlinear Programming Methods	47

TABLE OF CONTENTS (Continued)

		Page
D.	AN INDIRECT METHOD - *THE APPROXIMATE-FLOW TECHNIQUE* . . .	
	1. Introduction	48
	2. Approximation of Routed Flow	49
	3. Approximation of Flow prior to Routing	54
	4. Application of Generalized Duality Theory	59
	5. Discussion	62
V.	SUMMARY AND CONCLUSIONS	66
	REFERENCES	70
	APPENDIX - SUMMARY OF GENERALIZED DUALITY THEORY	73

I. INTRODUCTION

A. THE CONTROL PROBLEM

The pollution of bodies of water adjacent to urban centers, due to storm-produced overflows from combined sewer systems, is rapidly becoming a serious, nationwide problem [1]. In seeking methods of combatting this problem, attention has been focused on two areas: (i) improving the quality of the overflows, through sewer separation or reduced treatment processes that can handle large flow rates, and (ii) reduction of the magnitude of the overflows, by (a) somehow reducing storm inflows to the sewer system, (b) using storage capabilities within the sewers themselves (*ambient storage*), or (c) construction of additional storage facilities within the system (*auxiliary storage*) [1].

The use of (b) or (c) (or their appropriate combination) has arisen as a particularly attractive alternative, due to generally lower predicted costs and potentially greater effectiveness in dealing with the overflow problem. The U. S. Environmental Protection Agency is currently supporting a number of research and development studies in this area [15]. The goal is to utilize storage capabilities in such a way that flood peaks in the system can be lowered to a degree consistent with maximum advanced treatment plant inflow rates. Direct control is carried out through computerized remote operation of intake and outlet valves, regulators, adjustable weirs placed in sewers to effect ambient storage, etc. [2]. The complex and large-scale nature of the storage control problem should be readily apparent, since there may be hundreds of control points throughout the sewer system of a large urban center. There is critical need to take full advantage of current advances in computer technology (hardware and software) and

systems engineering. The high speed digital computer is required at all levels - from data collection and processing to implementation of sophisticated control logic.

Effective direct storage control, however, cannot be executed without intensive investigation in the following areas [7]:

1. storm prediction modeling
2. rainfall-runoff modeling
3. hydraulic modeling of the sewer system
4. design and operation of sensor networks for detecting rainfall and sewer flow rates
5. statistical analysis of noise-corrupted measurements and information, used as input for control strategies.

The above studies are necessary for accurately forecasting the magnitudes of flow rates throughout the sewer system, due to rainstorm activity. Ideally, this information is utilized in an automated *feedback control* process. As a storm passes over an urban center, sensors detect increasing rainfall and sewer flow rates. This information is passed to a computer control center via telecommunication and is fed into flow prediction models, from which a control strategy is generated, based on a programmed control logic. As the control strategy is generated and implemented, new information is detected as the storm continues, and the cycle continues, resulting in control strategies that are periodically monitored and updated in such a way as to effectively respond to the uniqueness of a particular storm event.

B. OFF-LINE VS. ON-LINE OPTIMIZATION

Our particular concern here is with control logic development, since the complexities involved seem to have impeded progress in this area. Some

of the difficulties to be expected are discussed by McPherson [22]. A number of cities such as Cleveland [2], Detroit [1], Seattle [17], Chicago [15], and San Francisco [26] appear to be in the early stages of control system development, but little specific information on control logic studies is currently available. Aside from some incomplete work by Bell, and others [4,5,6, and 7], which is summarized in a subsequent chapter of this report, accomplishments are meager in this area.

The enormous number of control alternatives possible precludes anything but application of modern systems techniques, particularly in the area of optimization theory. Control logic is determined through formulation of the control problem as an optimization problem where we seek to minimize total weighted overflows from the combined sewer system, subject to a number of constraints. The constraints include: (i) mass-balance equations describing the dynamics of flow and storage throughout the system, and (ii) physical limitations placed on flow rates and quantities in storage, due to: the dimensions of the sewers, capacities of ambient and auxiliary storage, and capacities of treatment plant facilities. The mass-balance equations are based on models constructed to simulate the behavior of the system. In general, realistic flow models result in complex optimization problems, so that studies are needed to determine the optimum trade-off.

There is question as to whether optimization should be carried out all *off-line*, all *on-line*, or a mixture of the two. Off-line optimization results in general operating policies, based on historical rainfall data, which are programmed into the control computer operating the system. On-line optimization, on the other hand, is carried out by the control computer in real time, and is based on historical records augmented by the particular storm occuring at the moment. It appears that a combination

of the two is necessary. Some on-line work is required, since it is impossible to model all possible storm situations in an off-line manner. It is, however, generally limited to simplified sewer flow models (e.g., linear), so that the optimization algorithm can be guaranteed to find a global solution. Off-line studies are free to use more realistic models, and therefore serve to augment the on-line work. The primary emphasis here is on the former.

C. SYSTEM DECOMPOSITION

The large-scale nature of the optimization problem requires that attempts be made to decompose the sewer system into a set of mildly interconnected subsystems or *subbasins*, which are temporarily disconnected. For example, the San Francisco system seems particularly well suited to decomposition, as schematically represented in Figure I-1 [26].

The *advantages* of decomposing a large-scale system are the following:

1. Greater conceptual understanding of the behavior of the system is attained when effort is made to identify and analyze subparts or subsystems within the large-scale system.
2. Mathematical programming techniques are available [21], such that interconnections between the subsystems can be temporarily cut, and control policies developed for the isolated subsystems. Each subsystem is then concerned with a limited number of control variables and a fraction of the total amount of data is necessary to operate the system. The result is considerable increase in system reliability toward achieving the overall system goals. The subsystems can then be recomposed together by a master control which achieves the recomposition in some kind of iterative fashion.

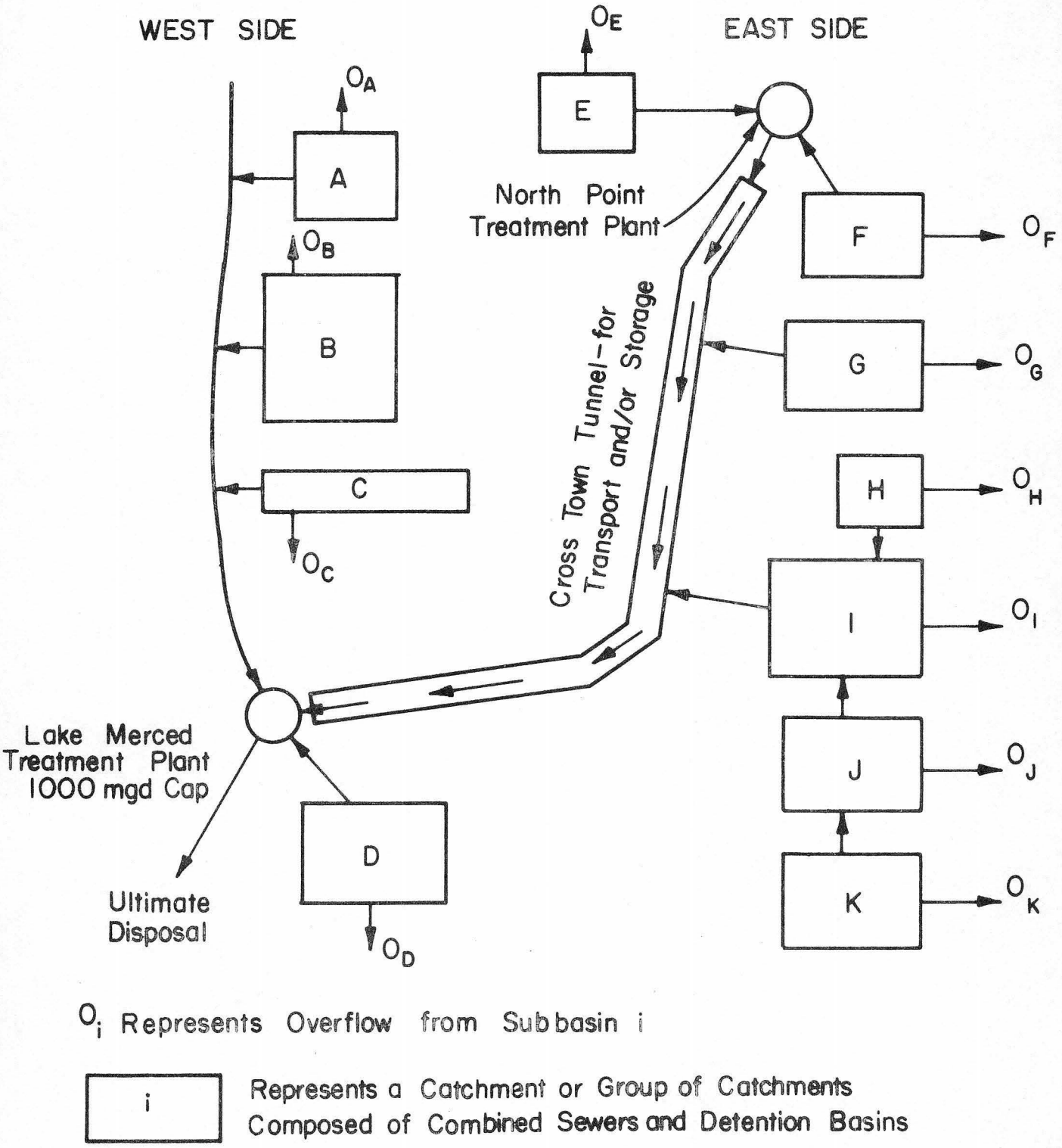

FIGURE I-1

DECOMPOSITION OF THE SAN FRANCISCO
COMBINED SEWER SYSTEM

3. Generally, less computer hardware is required for the decomposition approach than for centralized approaches. Essentially, computer storage is replaced by additional computer time. Less required computer hardware usually means greater reliability.

The emphasis in this report is on smaller scale subbasin analysis. Future reports will deal with development of master controllers that tie the subbasins together. With this plan in mind, we will use simplified storage configurations in discussing the various optimization formulations and solution strategies, thus preventing unwieldy notation in the presentation. Extensions to more complicated configurations should be reasonably obvious.

D. OBJECTIVES

The undertaking of this particular study has been motivated by the following:

1. The need for a broad, comprehensive evaluation of the basic optimization methodologies with regard to their specific applicability to solution of the optimal control problem for combined sewers.
2. The need for summarizing and critically analyzing current published attempts at formulating and solving the control problem via particular optimization techniques. As mentioned previously, however, little is available at the present time.
3. The necessity for generating new ideas with regard to specific optimization strategies for dealing with the complexities of the control problem that have so far hindered actual implementation for real time systems.

The basic objective here is to attempt to satisfy the above needs. Chapter II is concerned with analyzing which of the following broad categories is most applicable to our problem: finite or infinite-dimensional optimization methods. One should decide at an early stage which of these avenues to explore, before specific optimization strategies can be formulated. Current applications of infinite-dimensional optimization (or continuous-time optimal control theory) are considered in Chapter III, mainly based on the work of Bell, et. al. [4,5,6, and 7], and are critically evaluated. Chapter IV explores finite-dimensional techniques such as linear, nonlinear, and dynamic programming, and concludes with some ideas on application of indirect or dual approaches to the control problem. These approaches revolve around the concept of *approximate-flow*, and it is the author's opinion that they open the door to dealing with the difficulties that have so far hindered direct application of more conventional optimization techniques.

II. FINITE AND INFINITE-DIMENSIONAL OPTIMIZATION

A. A RESERVOIR CONTROL PROBLEM

A.1 Discrete Time Case [finite-dimensional optimization]

Suppose we are concerned with minimizing overflows at a particular control point i (≥ 2).

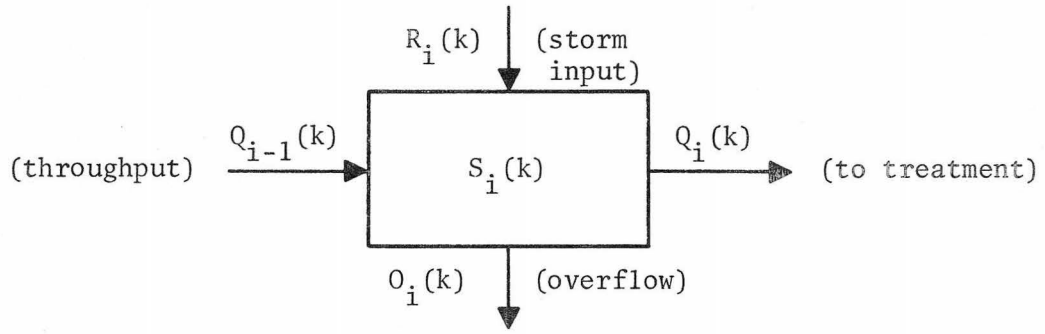

FIGURE 1
COMBINED SEWER STORAGE

where the time horizon is broken up into M discrete intervals

$$0 \triangleq t_1 < t_2 < \cdots < t_M < t_{M+1} \triangleq T_f$$

where interval k is defined by $[t_k, t_{k+1}]$, where $\Delta t = t_{k+1} - t_k$ for all $k = 1,\ldots,M$. For this problem

$S_i(k)$ = the storage (i.e., ambient and/or auxiliary) in the sewer at control point i, at the beginning of time period k (i.e., at time t_k)

$R_i(k)$ = the average rate of direct stormflow input to control point i, during time period k

$Q_i(k)$ = the average rate of throughput in the sewer from control point i, during period k

$O_i(k)$ = the average rate of overflow to receiving waters from control point i, during period k

$Q_{i-1}(k)$ = the sewer throughput rate from upstream control point $i - 1$, during period k.

Since our goal is to minimize overflows, an optimization problem can be formulated. In formulating this problem let us assume $Q_{i-1}(k)$ is given for all k, and temporarily drop the subscript i. Therefore, we can lump $Q_{i-1}(k)$ into the term $R(k)$ as given input to control point i

[Problem A1]: [the ω_k $(k = 1,\ldots,M)$ are weighting factors]

$$\underset{S(k),O(k),Q(k),\ k=1,\ldots,M}{\text{minimize}} \quad \sum_{k=1}^{M} \omega_k O(k)\Delta t \tag{1}$$

subject to:

dynamics (or state equation) $$S(k+1) = S(k) - [Q(k) + O(k) - R(k)]\Delta t \quad (k = 1,\ldots,M) \tag{1a}$$

initial condition $$S(1) = c \quad \text{(given)}$$

state-space constraint $$0 \leq S(k) \leq S_{max}, \quad k = 2,\ldots,M \tag{1c}$$

final condition $$S_{M+1} = S_{final} \quad \text{(may be specified)} \tag{1d}$$

control constraint $$0 \leq Q(k) \leq Q_{max}, \quad k = 1,\ldots,M \tag{1e}$$

where S_{max} and Q_{max} are upper bounds on storage and throughput, respectively. If $S(k)$ represents ambient storage, then S_{max} can be considered as a variable $S_{max}(k)$, where some kind of adjustable weir is utilized in the sewer. Then we would add the constraint

$$S_{max}(k) \leq \overline{S}_{max} \quad \text{for all } k$$

where $\overline{S}_{max}$ is the upper bound on storage obtained when the weir height is maximized.

Definitions

$S(k)$ $\overset{\Delta}{=}$ the state variable, or the state of the system at any time k. It is a dependent variable, since it is a function of $Q(k')$, $O(k')$, $k' = 1,\ldots,k-1$

$Q(k)$, $O(k)$ $\triangleq$ the *control* or *decision variables*, since they are independent variables and directly controllable

k $\triangleq$ the particular *stage* of the dynamic process.

Problem A1 is a straightforward linear programming problem. There are several other ways of formulating this single reservoir problem, but they involve introduction of some degree of nonlinearity. For example, suppose we let $\overline{Q}(k)$ represent total outflow from the reservoir (including overflows). The objective function then becomes

$$\min_{\substack{S(k),\overline{Q}(k),\\ k=1,\ldots,M}} \sum_{k\in K} [S(k) - S_{max}]$$

where (S_{max} no longer an upper bound on $S(k)$)

$$K = \{k \mid S(k) - S_{max} \geq 0\}$$

and the state equation is

$$S(k+1) = S(k) - \overline{Q}(k) + R(k), \quad k = 1,\ldots,M$$

Even though we now have only one decision variable $\overline{Q}(k)$, the objective function is piecewise linear, but not linear. This problem, however, is solveable by dynamic programming, which will be discussed further in a subsequent chapter.

A.2 Continuous-Time Case [infinite-dimensional optimization]

Suppose we let $\Delta t \to 0$, or equivalently, let $M \to \infty$. That is, Equation (1a) can be written as

$$\frac{S(t_k + \Delta t) - S(t_k)}{\Delta t} = -[Q(t_k) + O(t_k) - R(t_k)]$$

Taking the limit $\Delta t \to 0$ of both sides yields

$$\frac{dS(t)}{dt} = -[Q(t) + O(t) - R(t)]$$

(for all $0 \leq t \leq T_f$)

Therefore, the continuous-time version of Problem A1 is [Problem A2]:

$$\underset{S(t),O(t),Q(t),\ \text{for all } t\in[0,T_f]}{\text{minimize}} \quad \int_0^{T_f} \omega(t)O(t)\,dt \tag{2}$$

subject to:

dynamics (or state equation) $$\frac{dS(t)}{dt} = -[Q(t) + O(t) - R(t)], \quad t \in [0,T_f] \tag{2a}$$

initial condition $$S(0) = c \quad \text{(given)} \tag{2b}$$

state-space constraint $$0 \leq S(t) \leq S_{max}, \quad \text{for all} \quad t\in[0,T_f] \tag{2c}$$

final condition $$S(T_f) = S_{final} \quad \text{(may be specified)} \tag{2d}$$

control constraint $$0 \leq Q(t) \leq Q_{max}, \quad \text{for all} \quad t\in[0,T_f] \tag{2e}$$

A.3 Discussion

For the practical problem of optimally controlling combined sewer overflows via storage regulation, it is safe to assume that controls will be carried out in discrete time intervals. This is due to the following factors associated with on-line, automated control:

1. There is a finite amount of time required to actually effect control. That is, time is required for passage of information, the opening and closing of valves and regulators, the inflation and deflation of adjustable weirs, etc.
2. On-line control requires the processing of rainfall and sewer flow data, which is sampled at discrete-time [e.g., for the San Francisco system, data is collected every 15 seconds [26]].
3. Sufficient data must be collected in order to make a reasonable prediction of future storm input so that the next control can be effected. There is an interesting trade-off here:
   (a) Large intervals between control would allow the processing of more data, resulting in more accurate

prediction. Though the individual controls are more optimal in the sense that they are based on more accurate data, the system is less controllable due to the large intervals.

(b) Small intervals between control would result in less accurate storm prediction. Though the system is more controllable than in case (a), there is greater question as to the optimality of the controls.

Suppose it is decided that actual control of the system must occur between a discrete interval Δt_{actual} (which may be variable). Then there are two basic ways of determining the optimal controls $Q^*(k)$ and $O^*(k)$, where $\Delta t_{actual} = t_{k+1} - t_k$:

(i) Finite-Dimensional Optimization: Solve Problem A1, letting $\Delta t = \Delta t_{actual}/m$, where m is some integer ≥ 1, and determine $Q^*(k), O^*(k)$ from these results.

(ii) Infinite-Dimensional Optimization: Solve Problem A2, and determine $Q^*(t), O^*(t)$ for all $0 \leq t \leq T_f$, from which $Q^*(t_k)$ and $O^*(t_k)$ can be easily found for all k.

We are ultimately interested in considering the very general control problem involving many reservoirs in a complex of interraction. There is the need, then, to utilize realistic flow routing methods, which will unfortunately introduce nonlinearities into the state equation. In addressing ourselves to the general, complex control problems, we must decide which of these two solution approaches [(i) or (ii)] is most appropriate for the particular problem at hand. In attempting to answer this question, we will utilize a very general formulation of the control problem.

B. NECESSARY CONDITIONS FOR DISCRETE-TIME OPTIMAL CONTROL

Consider the following general control problem, letting $\Delta t = 1$

[Problem B]:

$$\min_{x,u} \sum_{k=1}^{M} f(x(k),u(k)) + \phi(x(M+1)) \tag{3}$$

[where $x = (x(1),\ldots,x(M+1))$, $u = (u(1),\ldots,u(M))$ and $\phi(\cdot)$ is an added term associated with the final state.]

subject to:

dynamics	$x(k+1) = x(k) + g(x(k),u(k))$ $(k = 1,\ldots,M)$	(3a)
initial condition	$x(1) = c$ (given)	(3b)
state-space constraint	$q(x(k)) \leq 0, \quad k = 1,\ldots,M$	(3c)
final condition	$p(x(M+1)) = 0$	(3d)
control constraint	$h(x(k),u(k)) \leq 0, \quad k = 1,\ldots,M$	(3e)

which is equivalent to Problem A1 if we define

$$u(k) \triangleq (Q(k),O(k))$$

$$x(k) \triangleq S(k)$$

$$f(\cdot,u(k)) \triangleq O(k)$$

$$\phi(\cdot) \triangleq 0$$

$$g(\cdot,u(k)) \triangleq Q(k) - O(k) + R(k)$$

$$p(x(M+1)) \triangleq S(M+1) - S_{final}$$

$$q(x(k)) \triangleq \underbrace{\begin{bmatrix} S(k) - S_{max} \\ -S(k) \end{bmatrix}}_{(a)} \quad \text{or} \quad \underbrace{S(k)[S(k) - S_{max}]}_{(b)}$$

$$h(\cdot,u(k)) \triangleq \begin{bmatrix} Q(k) - Q_{max} \\ -Q(k) \end{bmatrix} \quad \text{or} \quad Q(k)[Q(k) - Q_{max}]$$

{notice that (a) and (b) are exactly equivalent}

In general, then, $u(k), x(k), g(\cdot,\cdot), q(\cdot), p(\cdot)$, and $h(\cdot,\cdot)$ can themselves be vectors, for all k. For generality, let us specify that $u(k) \in E^m$, $x(k) \in E^n$, $g(\cdot,\cdot) \in E^n$, $q(\cdot) \in E^{2n}$ [for case (a)], $q(\cdot) \in E^n$ [for case (b)], $p(\cdot) \in E^n$, and $h(\cdot,\cdot) \in E^{\ell}$, for all k. For Problem A1, then, $m = 2$, $n = 1$, and $\ell = 2$.

If we assume that all functions are differentiable for all x,u, then the necessary conditions for an optimal solution to Problem B are the Kuhn-Tucker conditions [29].

The Lagrangian for Problem B is:

$$\begin{aligned} L(x,u,\lambda,r,\rho,\eta) = & \sum_{k=1}^{M} f(x(k),u(k)) + \phi_T(x(M+1)) \\ & + \sum_{k=1}^{M} \lambda(k)\cdot[x(k) - x(k+1) + g(x(k),u(k))] \\ & + \sum_{k=2}^{M} \gamma(k)\cdot q(x(k)) + \rho\cdot p(x(M+1)) \\ & + \sum_{k=1}^{M} \eta(k)\cdot h(x(k),u(k)) \end{aligned} \tag{4}$$

If x^*,u^* solves Problem B, then the following conditions must be satisfied:

(a) feasibility

$$x^*(k+1) = x^*(k) + g(x^*(k),u^*(k)) \quad (k = 1,\ldots,M) \tag{5}$$

$$q(x^*(k)) \leq 0, \quad k = 1,\ldots,M \tag{6}$$

$$p(x^*(M+1)) = 0 \tag{7}$$

$$h(x^*(k),u^*(k)) \leq 0, \quad k = 1,\ldots,M \tag{8}$$

and there exist Lagrange multipliers $\lambda^* \in E^{nM}$, $\gamma^* \in (E^{2n(M)})^+$, $\rho^* \in E^n$, and $\eta^* \in (E^{\ell M})^+$, such that

(b) complementary slackness

$$\sum_{k=1}^{M} \gamma^*(k)\cdot q(x^*(k)) = 0 \tag{9}$$

$$\sum_{k=1}^{M} \eta^*(k) \cdot h(x^*(k), u^*(k)) = 0 \tag{10}$$

(c) stationarity

$$\nabla_u L(x^*, u^*, \lambda^*, \gamma^*, \rho^*, \eta^*) = 0 \tag{11}$$

$$\nabla_x L(x^*, u^*, \lambda^*, \gamma^*, \rho^*, \eta^*) = 0 \tag{12}$$

Again, these conditions are only necessary for an optimal solution. That is, they may be satisified at points other than the global optimum (e.g., maxima, local minima, saddle-points, etc.). To guarantee that these conditions are both necessary and sufficient (i.e., their simultaneous solution will yield the global solution x^*, u^*), we must assume additionally that [29]:

(i) $f(\cdot,\cdot), \phi_T(\cdot)$ convex (or pseudo-convex)

(ii) $q(\cdot), h(\cdot,\cdot)$ convex (or quasi-convex)

(iii) $q(\cdot,\cdot), p(\cdot)$ linear

There is danger in using the Kuhn-Tucker conditions for finding x^*, u^* if these assumptions do not hold for a particular problem. In general, finite-dimensional optimization problems are not solved via the Kuhn-Tucker conditions, but rather, direct methods are utilized which generally can guarantee covergence to a local optimum, under certain mild conditions. Infinite-dimensional optimization problems, on the other hand, many times are solved using the continuous-time version of the Kuhn-Tucker conditions. Continuous-time problems with a high degree of nonlinearity can therefore present serious computational difficulties.

C. NECESSARY CONDITIONS FOR CONTINUOUS-TIME OPTIMAL CONTROL

For the discrete-time problem (Problem B), a nominal time increment of $\Delta t = 1$ was assumed. As we let $\Delta t \to 0$ (or $M \to \infty$), we obtain the continuous-time version of Problem B [Problem C]:

$$\min_{x(t),u(t)} \int_0^{T_f} f(x(t),u(t))dt + \phi(x(T_f)) \qquad (13)$$

for all $t\varepsilon[0,T_f]$

subject to:

dynamics (or state equation)	$\dot{x}(t) = g(x(t),u(t))$, for all $t\varepsilon[0,T_f]$	(13a)
initial condition	$x(0) = c$	(13b)
state-space constraint	$q(x(t)) \leq 0$, for all $t\varepsilon[0,T_f]$	(13c)
final condition	$p(x(T_f)) = 0$	(13d)
control constraint	$h(x(t),u(t))$, for all $t\varepsilon[0,T_f]$	(13e)

Since the Kuhn-Tucker conditions apply to the discrete-time problem for finite M, then the limiting conditions as $M \to \infty$ must be the necessary conditions for an optimal solution to Problem C.

Equation (11) is

$$\nabla_u f(x^*(k),u^*(k)) + \lambda^*(k)\cdot\nabla_u g(x^*(k),u^*(k)) + \eta^*(k)\cdot\nabla_u h(x^*(k),u^*(k)) = 0 \qquad (14)$$

(for $k = 1,\ldots,M$)

and Equation (12) is

$$\nabla_x f(x^*(k),u^*(k)) + \lambda^*(k) - \lambda^*(k-1) + \lambda^*(k)\cdot\nabla_x g(x^*(k),u^*(k)) + \gamma^*(k)\cdot\nabla_x q(x^*(k)) + \eta^*(k)\cdot\nabla_x h(x^*(k),u^*(k)) = 0 \qquad (15)$$

(for $k = 2,\ldots,M$)

$$\nabla_x \phi(x^*(M+1)) - \lambda^*(M) + \rho^*\cdot\nabla_x p(x^*(M+1)) = 0 \qquad (16)$$

C.1 From Discrete-Time to Continuous-Time

The continuous-time necessary conditions can be placed in a more concise format if we define the following function, called the *modified Hamiltonian* [25]†

† [pg. 110]

$$\bar{H}(x(k),u(k),\lambda(k),\gamma(k),\eta(k)) \triangleq H(x(k),u(k),\lambda(k)) + \gamma(k)\cdot q(x(k)) + \eta(k)\cdot h(x(k),u(k)) \tag{17}$$

where

$$H(x(k),u(k),\lambda(k)) \triangleq f(x(k),u(k)) + \lambda(k)\cdot g(x(k),u(k)) \tag{18}$$

and is called the *Hamiltonian*.

Therefore, we can replace Equation (13) with

$$\nabla_u \bar{H}(x^*(k),u^*(k),\lambda^*(k),\gamma^*(k),\eta^*(k)) = 0 \quad (k = 1,\ldots,M) \tag{19}$$

and Equations (14) and (15) with

$$\lambda^*(k-1) = \lambda^*(k) + \nabla_x \bar{H}(x^*(k),u^*(k),\lambda^*(k),\gamma^*(k),\eta^*(k)) \quad (k = 2,\ldots,M) \tag{20}$$

$$\lambda^*(M) = \nabla_x \phi(x^*(M+1)) + \rho^*\cdot\nabla_x p(x^*(M+1)) \tag{21}$$

Taking the limits of (18), (19) and (20)

(for all $k = 1,\ldots,M+1$) [28]†

$$\lim_{\Delta t \to 0} \nabla_u \bar{H}(x^*(k),u^*(k),\lambda^*(k),\gamma^*(k),\eta^*(k)) = \nabla_u \bar{H}(x^*(t),u^*(t),\lambda^*(t),\gamma^*(t),\eta^*(t)) \quad (\text{for all } t \in [0,T_f]) \tag{22}$$

$$\lim_{\Delta t \to 0} \frac{\lambda^*(k) - \lambda^*(k-1)}{\Delta t} = \lim_{\Delta t \to 0} \frac{\lambda^*(k+1) - \lambda^*(k)}{\Delta t} = \dot{\lambda}^*(t) = \nabla_x \bar{H}(x^*(t),u^*(t),\lambda^*(t),\gamma^*(t),\eta^*(t)) \quad (\text{for all } t \in [0,T_f]) \tag{23}$$

$$\lim_{\Delta t \to 0} \lambda^*(M) = \lambda^*(T_f) = \nabla_x \phi(x^*(T_f)) + \rho^*\cdot\nabla_x p(x^*(T_f)) \tag{24}$$

---

†[pgs. 447-8]

C.2 Necessary Conditions

We can now write the necessary conditions for the continuous-time problem. If $x^*(t) \in E^n$, $u^*(t) \in E^m$ (for all $t \in [0,T_f]$) solve Problem C, then the following conditions must be satisfied:

(a) feasibility

$$\dot{x}^*(t) = g(x^*(t),u^*(t)), \quad x(0) = c \tag{25}$$

$$q(x^*(t)) \leq 0 \tag{26}$$

$$p(x^*(T_f)) = 0 \tag{27}$$

$$h(x^*(t),u^*(t)) \leq 0 \tag{28}$$

and there exist multipliers $\lambda^*(t) \in E^n$, $\gamma^*(t) \in (E^{2n})^+$ (or $\in (E^n)^+$), $\rho^* \in E^n$, and $\eta^*(t) \in (E^\ell)^+$, such that

(b) complementary slackness

$$\gamma^*(t) \cdot q(x^*(t)) = 0 \tag{29}$$

$$\eta^*(t) \cdot h(x^*(t),u^*(t)) = 0 \tag{30}$$

(c) stationarity [from (22), (23) and (24)]

$$\nabla_u \overline{H}(x^*(t),u^*(t),\lambda^*(t),\gamma^*(t),\eta^*(t)) = 0 \tag{31}$$

$$\dot{\lambda}^*(t) = -\nabla_x \overline{H}(x^*(t),u^*(t),\lambda^*(t),\gamma^*(t),\eta^*(t)) \tag{32}$$

$$\lambda^*(T_f) = \nabla_x \phi(x^*(T_f)) + \rho^* \cdot \nabla_x p(x^*(T_f)) \tag{33}$$

These conditions correspond to the necessary conditions obtained from application of variational theory directly to Problem C [27]. Equation (32) is called the *adjoint equation* and Equation (33) is called the *transversality condition*.

C.3 Solution Difficulties

It appears on the surface that transformation from the discrete case to the continuous case is straightforward. A number of serious difficulties arise in the continuous case, however, that are not evident in the discrete case. Notice that the above necessary conditions are valid, in general,

only if $u^*(t)$ is continuous over $[0,T_f]$. The existence of the state-space $(q(\cdot) \leq 0)$ and control $(h(\cdot,\cdot) \leq 0)$ constraints, however, will tend to produce discontinuities in u^* at a finite number of points $\tau_j \in [0,T_f]$, $j = 1,\ldots,J$, where $\tau_1 < \tau_2 < \cdots \tau_J$. An additional set of necessary conditions, called *corner or jump conditions* [11] is therefore required for these τ_j. In general, the more state-space and control constraints there are, the more points of discontinuity there will be; and hence, the number of conditions to be satisfied increases proportionately.

One method of alleviating the problem of added corner conditions is to place (13c) and (13e) into the objective function (13) through use of arbitrary penalty functions. For example

$$P = \int_0^{T_f} f(x(t),u(t))dt + \phi(x(T_f)) + K_1 \int_0^{T_f} (q(x(t))^2 dt + K_2 \int_0^{T_f} (h(x(t),u(t))^2 dt \qquad (34)$$

is an example penalty function. The result is an unconstrained optimal control problem, where the above necessary conditions are applicable. The parameters K_1 and K_2 are adjusted until an optimal solution x^*,u^* is produced which satisfies the constraints. Also, a penalty of the form

$$K_3(p(x(T_f))^2$$

could be added also, allowing elimination of the transversality conditions. Penalty function methods, however, generally suffer from convergence problems, especially if the constraints can be satisfied only as $K_i \to \infty$, for all i. There are other difficulties associated with penalty function methods, as discussed in Bryson and Ho [9].

Even if it is possible to reduce the number of necessary conditions, there remains the difficult *two-point boundary-value problem* to be solved. That is, Equations (25) and (32) must be solved simultaneously, where there

are only given final conditions associated with (32) (the transversality conditions (33)). Since numerical integration methods require that all initial conditions be given, $\lambda(0)$ must be guessed and adjusted until the transversality condition is satisfied. This trial and error procedure is computationally inefficient, and if a high degree of nonlinearity exists in the problem, the current solution may never be attained. This is primarily due to *instability* difficulties, where small changes in $\lambda(0)$ produce large changes in $\lambda(T_f)$.

Solution procedure generally starts with determination of $u^*(x(t),\lambda(t),t)$ as a function of x and λ from Equation (31), and then the two-point boundary-value problem is attempted. An alternative to using (31) is to apply the *Maximum Principle* [25]. Here, the Hamiltonian H is utilized in place of $\overline{H}$ (so that the complementary slackness conditions can be eliminated) and (31) is replaced with

$$H(x(t),u^*(t),\lambda(t)) \leq H(x(t),u(t),\lambda(t)) \tag{35}$$

for all u satisfying (28), from which $u^*(x(t),\lambda(t),t)$ can hopefully be determined. It may not be possible to determine $u^*(x(t),\lambda(t),t)$ from (31) or (35). The so-called *singular* case is an example, where control $u(\cdot)$ appears linearly in $f(\cdot,\cdot)$ (or not at all), and so can not be explicitly determined from (31).

We see, then, that aside from the inherent dangers of using necessary conditions to find x^*,u^*, the actual solution can be extremely difficult for nontrivial continuous control problems. Considerable effort has therefore been directed, in recent years, towards applying methods original developed for finite-dimensional problems to infinite-dimensional problems. These would be termed "direct" methods, since the necessary conditions are essentially ignored and an initial guess $x^0(t),u^0(t)$ (for all $t \in [0,T_f]$) starts an iterative process that attempts to successively decrease the

objective function (subject to the constraints) in some fashion. Some of the methods that have been applied include gradient methods, conjugate directions, Newton-Raphson, and others. In general, these methods are more difficult to apply to infinite-dimensional problems than finite-dimensional problems, with computer storage capabilities being a consistent constraint.

D. SUMMARY AND DISCUSSION

Let us summarize what has been shown in this chapter:

1. There are two basic approaches to solving the optimal control problem of minimizing overflows from combined sewer systems:
   (a) Solve the finite-dimensional problem [Problem B], where the time horizon has been discretized, and determine the optimal controls for each interval.
   (b) Solve the infinite-dimensional problem [Problem C] and discretize the resulting continuous-time optimal control[†] according to the interval Δt_{actual}
2. The necessary conditions for the continuous-time optimal control problem can be derived as limiting versions (as $\Delta t \to 0$) of the Kuhn-Tucker necessary conditions for the discrete-time problem.
3. Infinite-dimensional optimization is more heavily dependent upon utilizing necessary conditions for determining optimal controls than is finite-dimensional optimization. Since necessary conditions are generally applicable at local minima, maxima, saddle-points, etc., solution results can be deceiving for nonlinear problems (unless conditions (i) - (iii) of Section B hold, thus assuring that the Kuhn-Tucker conditions are both necessary and sufficient).

[†]Note: Since integration must be carried out numerically on a digital computer, then this control will actually be discretized, though the time intervals used for integration $\delta t << \Delta t_{actual}$.

4. The necessary conditions for infinite-dimensional problems are difficult to solve simultaneously (for x^*, u^*) because:
   (a) Large numbers of constraints (on control and state variables) tend to create large numbers of necessary conditions (corner conditions) and the control logic becomes increasingly complex.
   (b) Computational inefficiency arises in solution of the two-point boundary-value problem, and the possibility of divergence is ever-present for nonlinear problems, due to instability.
5. Data for the combined sewer problem are taken in discrete-time. But notice, for example, that Problem A2 requires that continuous data $R(t)$ (for all $t \in [0,T_f]$) be given. Thus, a continuous curve must be approximated from the discrete data. Since there are an infinite number of such approximations (based on whatever fitting criteria are used), the uniqueness of the resulting optimal control $u^*(t)$ may be questionable.

These statements seem to suggest that finite-dimensional optimization is superior, at least for our problem. Notice, however, that if M is large (which may be necessary for accurate control), that the number of variables involved in Problem B would quickly tax the rapid-access storage capacity of even the largest digital computers. If this is the case, there may be no other alternative but to apply continuous-time control theory. On the other hand, we could arbitrarily decrease M (i.e., increase Δt) so that Problem B becomes solveable, with a resulting decrease in the accuracy of the control. Though the resulting u^* is optimal with respect to these coarser intervals, it will probably be suboptimal with respect to the more realistic finer intervals.

For the combined sewer problem, it appears that M can be kept to a reasonable size (allowing solution by finite-dimensional methods), due to statements 1, 2, and 3 in section A.3 of this chapter. In addition, control policies can probably be developed storm by storm, so that a problem need not be defined over several storms. As Canon, et. al. [10],† have succinctly stated, the "...main reason for attaching so much importance to discrete optimal control is technical and stems from the constantly increasing use of digital computers in the control of dynamical systems. In any computation carried out on a digital computer, we can do no better than obtain a finite set of real numbers. Thus, in solving a continuous optimal control problem... we are forced to resort to some form of discretization." The question, then, is whether to discretize prior to computation (as in finite-dimensional optimization) or during and subsequent to computation (as in infinite-dimensional optimization). For the combined sewer problem, the author's recommendation is that the former be stressed.

The following chapter will serve to support the above conclusions concerning infinite-dimensional optimization, as it is applied to some simplified subbasin configurations. This will be followed by a chapter on finite-dimensional optimization techniques, concluding with a proposed solution procedure based on recent advances in duality theory.

---

†[pgs. 1 and 2]

III. APPLICATIONS OF CONTINUOUS-TIME CONTROL THEORY

A. INTRODUCTION

The following applications represent the first attempts at solving the optimal control problem for combined sewer systems, as far as this author is aware. Most of the work has been carried out by W. Bell, and reported in [5], [6], and [7]. It may be valuable to attempt solution of a control problem by continuous-time theory before finite-dimensional work, since finite-dimensional optimization tends to require a larger initial investment in computer time and demands a greater quantity of computer storage. For nonlinear, nonconvex problems, it is usually impossible to determine *a priori* whether or not these attempts will be successful. As it becomes increasingly evident that the difficulties are insurmountable, effort should be shifted to finite-dimensional optimization.

Such is the experience with the combined sewer control problem. Considerable difficulty has been encountered with applying continuous-time theory to even very idealized subbasin configurations of at most two or three reservoirs, with time lag in flow routing neglected. This experience has discouraged further extension to more realistic configurations, and current research effort is concentrating on solving the control problem by finite-dimensional optimization techniques. However, the initial efforts in continuous-time control are reported here for the following reasons:

1. To give evidence as to the viability of shifting emphasis to finite-dimensional optimization.
2. Limited results have been obtained for certain specialized cases, and it is hoped that they will serve to aid in generating accurate initial guesses for direct solution of the finite-dimensional problems.

B. AMBIENT STORAGE

B.1 Two Reservoirs in Series

Bell and Wynn [5] have applied continuous-time control theory to the problem of minimizing overflows from a system composed of two reservoirs in series, where storage is defined in terms of water accumulation behind an adjustable weir placed in the sewer.

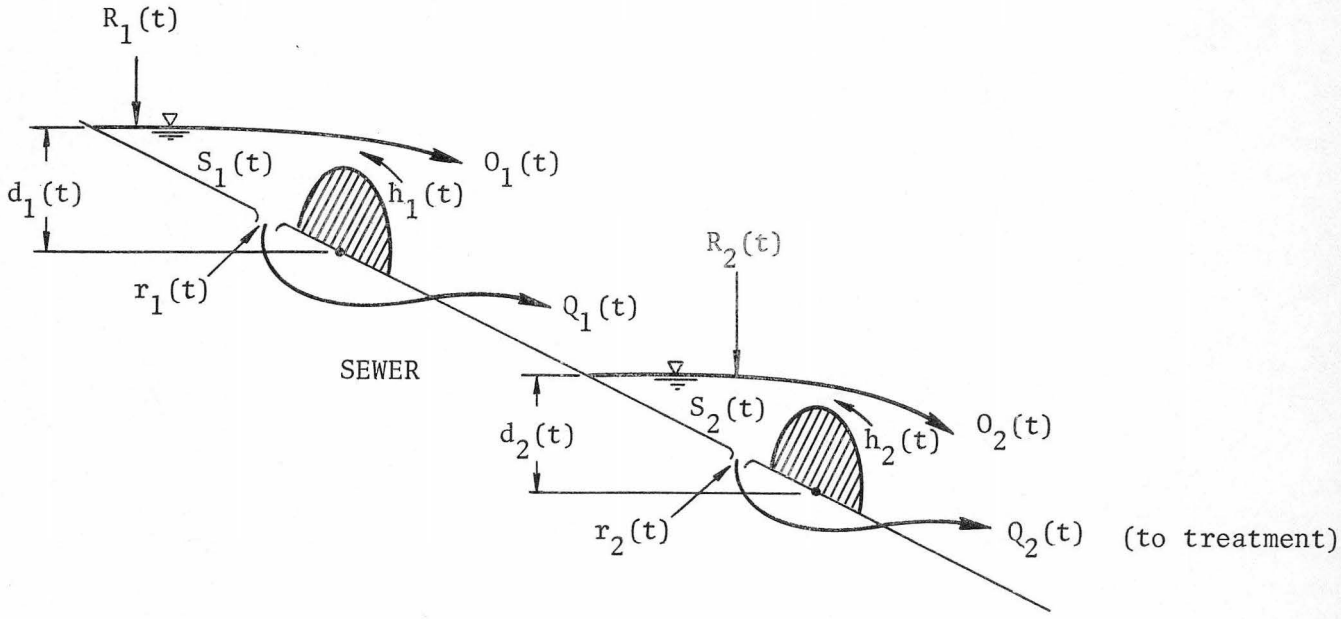

FIGURE III-1
AMBIENT STORAGE IN SERIES

where (at time t)

$R_i(t)$ = rate of direct input to storage behind weir i

$S_i(t)$ = accumulated storage behind weir i

$O_i(t)$ = rate of overflow from reservoir i to receiving waters

$Q_1(t)$ = throughput rate from reservoir 1 to reservoir 2

$Q_2(t)$ = rate of flow to treatment plant

$r_1(t)$ = variable radius of orifice controlling throughput from reservoir 1

$r_2(t)$ = variable orifice radius for flow to treatment

$h_i(t)$ = the head over weir i

$d_i(t)$ = depth of flow at weir i

In the example one-reservoir problem discussed in Chapter II [Problem A2], $Q(t)$ and $O(t)$ represented the control variables, and $S(t)$ was the state variable. The state variable $S(t)$ is actually a direct, one-to-one function of $d(t)$, so we can replace $S(t)$ with $d(t)$ as the state variable. For flow through an orifice

$$Q_i(t) = a_i r_1(t)^2 \sqrt{d_i(t)}, \qquad i = 1,2 \tag{1}$$

where a_i is a given constant. Therefore, we can replace $Q_i(t)$ with $r_i(t)$ as a control variable. For flow over a weir

$$O_i(t) = b_i h_i(t)^{3/2}, \qquad i = 1,2 \tag{2}$$

where b_i is a given constant. Likewise, we replace $O_i(t)$ with $h_i(t)$ as a control variable. [Note: weir height is uniquely defined by h and d].

The following state equations are expressed with the assumption that there are negligible time lags. Future work should consider not only time lags, but backwater effects in the flow routing. By conservation of mass

$$\dot{d}_1(t) = [R_1(t) - a_1 h_1(t)^{3/2} - b_1 r_1(t)^2 \sqrt{d_1(t)}] \,/\, A_1(d_1(t)) \tag{3}$$

$$\dot{d}_2(t) = [R_2(t) + b_1 r_1(t)^2 \sqrt{d_1(t)} - a_2 h_2(t)^{3/2} - b_2 r_2(t)^2 \sqrt{d_2(t)}] \,/\, A_2(d_2(t)) \tag{4}$$

where $A_i(d_i(t))$, $i = 1,2$, are given area-depth relationships, characteristic of the sewer and its slope, such that

$$\dot{S}_i(t) = \dot{d}_i(t)A_i(d_i(t)), \qquad i = 1,2$$

Since we wish to minimize total volume of accumulated overflow, let $\dot{o}(t) = O(t)$, or

$$\dot{o}_i(t) = b_i h_i(t)^{3/2}, \qquad i = 1,2 \tag{5}$$

Hence, the $o_i(t)$ are introduced as additional state variables, and the infinite-dimensional problem is [weighting factors are represented by the vector $\omega = (\omega_1,\omega_2)$]:

$$\min_{d(t),o(t),r(t),h(t) \text{ for all } t \in [0,T_f]} \omega \cdot o(T_f) \tag{6}$$

[where $d(\cdot),o(\cdot),r(\cdot),h(\cdot) \in (E^2)^+$]

subject to:

state equations [Equations (3), (4), and (5)]

initial conditions $d(0),o(0)$ (given)

final conditions [none]

state-space constraints
$$d(t)\cdot[d(t) - d_{max}] \leq 0 \tag{7}$$

control constraints
$$\begin{cases} r(t)\cdot[r(t) - r_{max}] \leq 0 & (8) \\ h(t)\cdot[h(t) - d(t)] \leq 0 & (9) \\ R_2(t) + b_1 r_1(t)^2\sqrt{d_1(t)} \leq Q_{1,max} & (10) \end{cases}$$

where d_{max}, r_{max}, and $Q_{1,max}$ are given upper bounds. Notice that Inequalities (7), (8) and (9) are placed in the form of case (b) [Chapter II, Section B]. Inequality (9) assures that $h(t)$ cannot

exceed $d(t)$, whereas (10) states that direct input to reservoir 2, plus throughput from reservoir 1, must be less than or equal to the maximum capacity of the sewer $Q_{1,max}$ between reservoirs 1 and 2.

In terms of the general format of Problem C [Chapter II, Section C]

$$u(t) \triangleq (r(t),h(t)) \Rightarrow m = 4$$

$$x(t) \triangleq (d(t),o(t)) \Rightarrow n = 4$$

$$f(\cdot,\cdot) \triangleq 0$$

$$\phi(x(T_f)) \triangleq \omega \cdot o(T_f)$$

...and so on. The necessary conditions [Chapter II, Section C.2] are now written, including appropriate corner conditions, and an attempt is made to find $r^*(t),h^*(t),d^*(t),o^*(t)$ that will satisfy them.

Bell has written a computer program for this problem and discussed the results in an unpublished report [4]. To avoid the difficult two-point boundary-value problem, a successive approximation scheme is utilized, which allows the state equations to be solved independently of the adjoint equations. Referring to the general format of Problem C, Chapter II, the procedure is basically:

(a) Guess an initial feasible estimate of $u^* \triangleq \{u^*(t), 0 \leq t \leq T_f\}$, and call it u^o. Set iteration number $j = 0$.

(b) Solve the state equations by numerical integration and determine feasible $x^{j+1} \triangleq \{x^{j+1}(t), 0 \leq t \leq T_f\}$.

(c) The given control u^j is now ignored and x^{j+1} is used in the simultaneous solution of the adjoint equations (II-32), the control equations (II-31), and the complementary slackness

conditions (II-29) and (II-30), yielding $u^{j+1}, \lambda^{j+1}, \gamma^{j+1}$, and η^{j+1}. [Note: the adjoint equations must be integrated backwards from $t = T_f$ to $t = 0$].

(d) At this point, all necessary conditions are satisfied, except for the fact that, in general, $u^j(t) \neq u^{j+1}(t)$, for all t. For some tolerance ε, an example "stop" criterion, not necessarily used by Bell, is:

$$|u^j(t) - u^{j+1}(t)| < \varepsilon, \quad \text{for all } t?$$

(i) If YES, STOP.

(ii) If NO, set $j \to j+1$; GO TO (b)

The results from experience with the computer program can be summarized as follows [4]:

1. The following assumptions were made:
   (i) $\omega_1 < \omega_2$, or greater weight was placed on overflows from the downstream reservoir
   (ii) $r_2(t)$ was eliminated as a variable, and assumed to be a given constant
   (iii) When $d_2(t) = d_{max}$, $h_2(t) > 0$. This simplifies the corner conditions, since we are excluding the possibility that $h_2(t) = 0$ when $d_2(t) = d_{max}$.
2. The computer program did not converge to the optimal solution, but tended to oscillate closely around it. The necessary conditions were analyzed by hand, and it was found that control tended to be of a *bang-bang* nature (i.e., instantaneous switching

occurred at various points, where control was transferred from one control constraint boundary to another). In general, the optimal control stayed on either state-space or control constraint boundaries at all times.

3. Singularity (as discussed in Chapter II) appeared as a consistent difficulty, since the objective function (6) is only defined at T_f. Therefore, at those times t where a certain number of the multipliers $\lambda(t), \gamma(t)$, or $\eta(t)$ vanish, then there is the possibility that one or more of the control variables will vanish from the control equations (II-31). The problem, then, is how to find a unique $u^*(t)$.

Bell [4] also examined a problem of two reservoirs in parallel. No computer program was written, but analysis of the necessary conditions by hand produced some approximate results. It was noted that control, again, was of the bang-bang type. Bell also found that the problem of singularity did not occur if the throughputs Q_i were placed in the objective function, along with appropriate weighting factors. Notice that there are two ways that this can be done. To the objective function, add either the term

$$\text{(i)} \quad -\int_0^{T_f} \left[\sum_{i=1}^{2} \nu_i(t) a_i r_i(t)^2 \sqrt{d_i(t)}\right] dt$$

where $\nu_i(t)$ are weighting factors, or

(ii) let $\dot{q}(t) = Q(t)$, and add the term $-[\nu \cdot q(T_f)]$.

It is not clear from [4] which approach Bell applied, but it seems that the danger of singularity still remains with approach (ii) [Notice the

minus sign, since we wish to maximize throughput]. It appears that approach (i) would eliminate the possibility of singularity, at least for $r(t)$.

In general, it must be concluded that the control logic for even these very simple problems, with several idealized assumptions, is complex enough to discourage further extension to more realistic problems. For this problem, the difficult two-point boundary-value problem was avoided by using a successive approximation scheme. For nonlinear problems, however, successive approximation methods can be highly unstable, even when initial guesses $u^o(t)$ are very close to $u^*(t)$ [24]. The lack of convergence here seems to present some evidence to this effect.

B.2 A Three-Reservoir Problem

Results from the two-reservoir problems previously described have been extended by Bell, Wynn, and Smith [7] to a three reservoir problem, whose configuration is shown in Figure III-2.

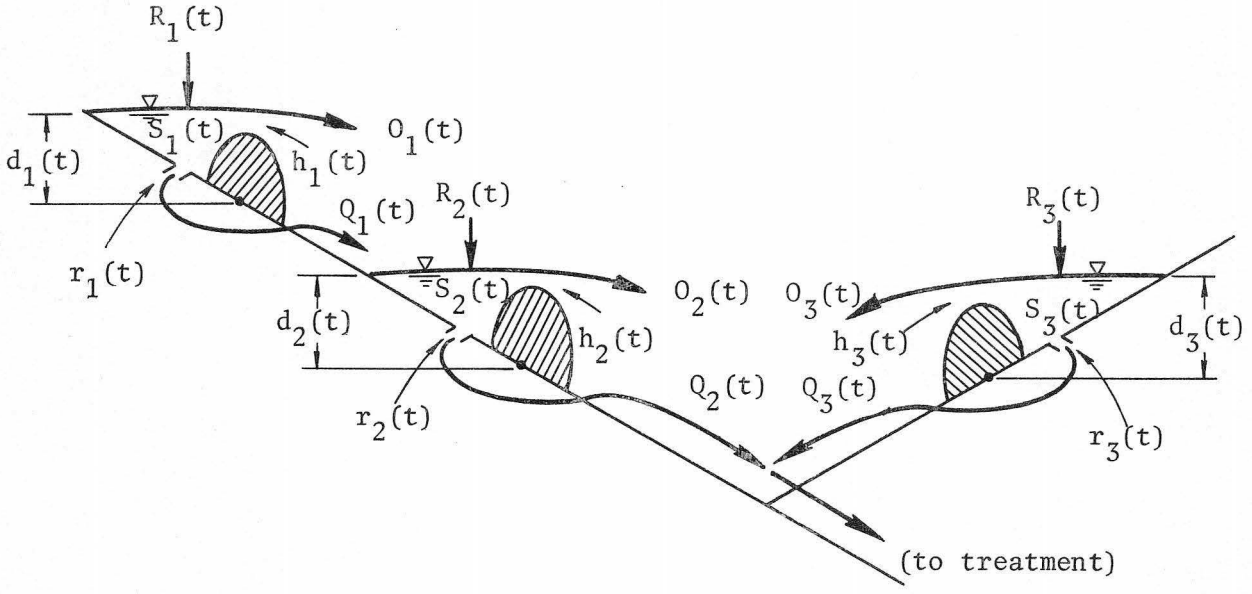

FIGURE III-2

A THREE-RESERVOIR PROBLEM FOR AMBIENT STORAGE

In extending the formulations for the two-reservoir problems to this three-reservoir problem, it is desired to introduce throughput into the objective function, along with appropriate weighting factors, so that the problem of singular control, hopefully, does not appear. Therefore, let $\dot{q}(t) = Q(t)$, or

$$\dot{q}_i(t) = a_i r_i(t)^2 \sqrt{d_i(t)}, \qquad i = 1,2,3 \tag{11}$$

so that $q(t)$ is considered as a new state variable.

The optimization problem is:

$$\min_{\substack{d(t),o(t),q(t),r(t),h(t) \\ \text{for all } t \in [0,T_f]}} [\omega \cdot o(T_f) - \nu \cdot q(T_f)] \tag{12}$$

[where $d(\cdot),o(\cdot),q(\cdot),r(\cdot),h(\cdot),\omega,\nu \in (E^3)^+$, and ν is the vector of weighting factors associated with accumulated overflows]

subject to:

state equations:

- [Equations (3) and (4)], plus
- $\dot{d}_3(t) = [R_3(t) - b_3 r_3(t)^2 \sqrt{d_3(t)} - a_3 h_3(t)^{3/2}] / A_3(d_3(t))$ (13)
- [Equation (5)], for i = 1,2,3
- [Equation (11)], for i = 1,2,3

initial conditions: $d(0),o(0)$, and $q(0)$ (given)

final conditions: [none]

state-space constraints:

$$d(t) \cdot [d(t) - d_{max}] \leq 0 \tag{14}$$

control constraints:

$$r(t) \cdot [r(t) - r_{max}] \leq 0 \tag{15}$$

$$h(t) \cdot [h(t) - d(t)] \leq 0 \tag{16}$$

$$R_2(t) + b_1 r_1(t)^2 \sqrt{d_1(t)} \leq Q_{1,max} \tag{17}$$

$$b_2 r_2(t)^2 \sqrt{d_2(t)} + b_3 r_3(t)^2 \sqrt{d_3(t)} \leq Q_{2,max} \tag{18}$$

Inequality (18) has been added to insure that total flow to the treatment plant does not exceed $Q_{2,max}$. In terms of the general format of Problem C

$$u(t) \triangleq (r(t),h(t)) \quad => \quad m = 6$$

$$x(t) \triangleq (d(t),o(t),q(t)) \quad => \quad n = 9$$

$$f(\cdot,\cdot) \triangleq 0$$

$$\phi(x(T_f)) \triangleq \omega \cdot o(T_f) - \nu \cdot q(T_f)$$

The objective function has been formulated in such a way that overflows are minimized and throughputs maximized, based on the choice of weighting factors. Selection of appropriate weighting factors will probably be based primarily upon the relative levels of pollution at the various control points. It appears that pollution would tend to increase downstream, so that $\omega_1 < \omega_2$ and $\nu_1 > \nu_2$. Again, time delays are neglected for this problem, as reflected in the state equations.

As before, the necessary conditions, including all corner conditions, are written for this problem, and an attempt is made to solve them simultaneously. The computer programs were developed for this purpose, and discussed in [4] and [7]:

1. The first program followed the same basic successive approximation scheme as carried out for the two reservoir problem, except that a perturbation procedure was included as an attempt to get around the problem of singularity. As far as this author is aware, no convergence has been attained as yet. If instability of the successive approximation method was a factor in the lack of convergence for the two-reservoir problem, then it seems that this would be further accentuated for the more complicated three-reservoir problem.

2. In order to simplify the control logic and reduce the total number of necessary conditions, a penalty function approach was utilized as an alternative. As discussed in Chapter II, penalty terms of the form seen in Equation (II-34) can be added for all the state-space and control constraints, thus leaving an unconstrained control problem. The problems associated with large numbers of corner conditions are avoided. As discussed in Chapter II, however, penalty function methods suffer from difficulties of their own. There is a large element of trial and error involved, and the control problem must be solved several times, each time adjusting the parameters K_1 and K_2, until the correct K_1^*, K_2^* are found such that the state-space and control constraints are indirectly satisfied. Bell [4] indicates that convergence has been attained in some cases, but that the following difficulties arose in connection with the penalty function method:

   (a) A saddle-point tended to occur at $r = 0$. Since we are dealing with necessary conditions, there was danger that a saddle-point solution would result, rather than the true global minimum. This was dealt with by requiring that $r \geq \varepsilon$, where ε is an arbitrarily small number.

   (b) Singularity still tended to be a problem, even though the addition of the penalty terms insures that the control variables appear explicitly in the control equations (as long as $K_i \neq 0$). Various attempts have been made to overcome this problem, but success is not assured as yet.

C. AUXILIARY STORAGE

For situations where auxiliary storage dominates over ambient storage (e.g., the San Francisco system), extensions of Problem A2 [Chapter II] are more appropriate. That is, instead of defining control variables in terms valve of settings and adjustable weir heights, we return to defining control in the broader sense of flow rates, as before. This may also be a viable alternative for analyzing ambient storage by continuous-time theory. The following formulation, therefore, is applicable to both ambient and auxiliary cases. Suppose we have been able to determine $Q_i^*(t)$, for all $t \in [0,T_f]$, as well as the optimal storage $S_i^*(t)$. It is a relatively simple matter to determine the optimal orifice settings $r_i^*(t)$ and head over the weir $h_i^*(t)$ from these values. Let us formalize these ideas by setting up the preceding three-reservoir problem as an extension of Problem A2:

$$\min_{\substack{S(t),O(t),Q(t) \\ \text{for all } t \in [0,T_f]}} \int_0^{T_f} [\omega(t)\cdot O(t) + \nu(t)\cdot Q(t)]dt$$

subject to:

state equations

$$\begin{cases} \dfrac{dS_1}{dt} = R_1(t) - O_1(t) - Q_1(t) \\ \dfrac{dS_2}{dt} = R_2(t) - O_2(t) + Q_1(t) - Q_2(t) \\ \dfrac{dS_3}{dt} = R_3(t) - O_3(t) - Q_3(t) \end{cases}$$

initial conditions $S(0)$ given

state-space constraints $0 \leq S(t) \leq S_{max}$

$$\text{control constraints} \begin{cases} O(t) \geq 0 \\ 0 \leq Q(t) \leq Q_{max} \\ Q_2(t) + Q_3(t) \leq \overline{Q}_{max} \end{cases}$$

where $S(\cdot), O(\cdot), Q(\cdot), \omega(\cdot), \nu(\cdot) \in (E^3)^+$, and we are using case (a) [Chapter II, Section B] for the control constraints to keep the problem linear.†

An alternative formulation, perhaps more amenable to efficient computation, and comparable to the approach used in the ambient storage case, is to define the objective function as

$$\min_{\substack{S(t),o(t),q(t),O(t),Q(t) \\ \text{for all } t \in [0,T]}} \omega \cdot o(T_f) + \nu \cdot q(T_f)$$

and add on the additional state equations

$$\dot{q}(t) = Q(t), \quad Q(0) \quad \text{given}$$

$$\dot{o}(t) = O(t), \quad O(0) \quad \text{given}$$

The above problem is a linear control problem, for which there are a number of highly developed, efficient algorithms for solving them [19]. These algorithms are based on applications of generalized linear programming [14], rather than standard optimal control theory, as described in Chapter II. These methods appear to have great potential for solving subbasin problems involving several more reservoirs than the simple three-reservoir example discussed here.

In order to apply standard optimal control theory, the objective function should be at least quadratic, so as to avoid the singularity problem discussed in Chapter II. One possibility would be to use a criterion of the form

---

†Notice that Q_{max} and $\overline{Q}_{max}$ are actually variables, since they are functions of S. For the above formulation, average values are used.

$$\omega \cdot (o(T_f))^2 + \nu \cdot (q(T_f))^2$$

though there is some question as to the equivalence of this criterion to our basic objective of minimizing overflows. The resulting problem has a quadratic criterion and a linear set of constraints. Perhaps no other subject in the area of continuous-time control theory has received more attention in past years than the linear-quadratic problem. The two-point boundary-value problem for this formulation is easily solved by the *sweep method* and solution of the well-known *matrix-Ricatti* equations [25].†

Computational experience is not yet available on application of the above ideas to the combined sewer problems. It appears, from the above discussion, that there would be less difficulty in applying these approaches to the ambient storage case, than that originally attempted by Bell [4].

D. DISCUSSION

There is little doubt that continued effort towards obtaining convergence for the ambient storage formulations will eventually succeed, particularly via penalty function approaches. As seen in some of the discussion concerning the three-reservoir problem, however, the effort tends to involve a good deal of problem manipulation, intuitive insight, and a measure of good luck. Again the major point we are emphasizing here is not the impossibility of solving individual control problems, but the great deal of effort involved in obtaining a solution. There seems, in general, to be a high level of programming skill required. Since the eventual hope is to consider models

---

†[pgs. 102-3]

1. composed of several interconnected reservoirs in a variety of configurations
2. which include realistic flow routing components that properly allow for time lag between control points
3. that consider backwater effects in flow routing. The ideal situation is that the full St. Venant equations are utilized,

so that from the above experience, it must be concluded that each particular model situation would require a unique effort in obtaining solutions that would probably not be applicable to other models.

The auxiliary storage formulation of the previous section was seen to also be applicable to the ambient storage case, and perhaps a more effective approach. It was shown that highly effective computational techniques are available for solving these problems, both for the completely linear case and the linear-quadratic case.

The major concern, however, is extension to more realistic flow routing techniques, thus introducing nonlinearities into the state equations. The highly efficient methods previously alluded to must then be abandoned, unless we attempt to linearize the nonlinear equations and converge to the solution of the original nonlinear problem in some kind of iterative fashion. This is the essence of a method called *Quasilinearization*, developed by Bellman and Kalaba [8]. Due to the nonconvexity of control problems with nonlinear state equations (i.e., nonlinear equality constraints result in a nonconvex constraint region, in general), this approach tends to be rather unstable [24]. If we attempt to return to application of standard continuous control theory to the nonlinear control problem, we are again thwarted by the difficult two-point boundary-value problem, as well as other hinderances, as discussed in Chapter II.

IV. MATHEMATICAL PROGRAMMING APPROACHES

A. INTRODUCTION

As was shown in Chapter II, finite-dimensional optimization appears to be better suited to the problem of optimal control of combined sewer overflow. We found this to be due particularly to:

(i) the physical nature of the system
(i.e., control is actually effected in discrete-time)

(ii) the difficulty of applying optimal control theory (infinite-dimensional optimization), since it is based on necessary conditions for optimality

Methods used to solve finite-dimensional optimization problems are lumped under the term *mathematical programming*. That is, linear, nonlinear, and dynamic programming are all mathematical programming techniques. The variety of techniques is large, particularly under the category of nonlinear programming. Again, mathematical programming methods usually are not based upon solution of necessary conditions, as in continuous-time control theory. Necessary conditions may be used, however, for checking the optimality of solutions determined by other means.

The purpose of this chapter is to discuss some of the techniques available and conclude with a methodology the author feels is most conducive to the problem at hand. Emphasis will be placed on the advantages and disadvantages of each technique, based on the following questions:

1. How realistic a model concerning the flow dynamics of the system can be utilized?
2. Can the method tolerate a large number of variables? That is, is it conducive to decomposition, since the large-scale problem must eventually be dealt with?

3. Will the method guarantee convergence to global or just local solutions?

The general finite-dimensional problem is repeated from Chapter II [Problem B]:

$$\min_{x,u} \sum_{k=1}^{M} f(x(k),u(k)) + \phi(x(M+1)) \tag{1}$$

[where $x = (x(1),\ldots,x(M+1))$, $u = (u(1),\ldots,u(M))$]

subject to:

$$x(k+1) = x(k) + g(x(k),u(k)) \quad (k = 1,\ldots,M) \tag{1a}$$

$$x(1) = c \tag{1b}$$

$$q(x(k)) \leq 0, \quad k = 1,\ldots,M \tag{1c}$$

$$p(x(M+1)) = 0 \tag{1d}$$

$$h(x(k),u(k)) \leq 0, \quad k = 1,\ldots,M \tag{1e}$$

The particular technique to be applied depends upon:

(i) the nature of $f(\cdot,\cdot)$ and $\phi(\cdot)$ (i.e., their linearity, nonlinearity, nonconvexity, continuity, etc.)

(ii) the nature of $g(\cdot,\cdot)$, $q(\cdot)$, $p(\cdot)$, and $h(\cdot,\cdot)$

(iii) the number of state variables (n) and decision or control variables (m) at each stage

B. AN EXAMPLE THREE-RESERVOIR PROBLEM

As explained in Chapter I, we are primarily interested in subbasin analysis here. Future work will concentrate on fitting the subsystems into a large-scale framework. Let us then consider an example subbasin configuration composed of three auxiliary reservoirs in series, with overflow possible from each reservoir, where

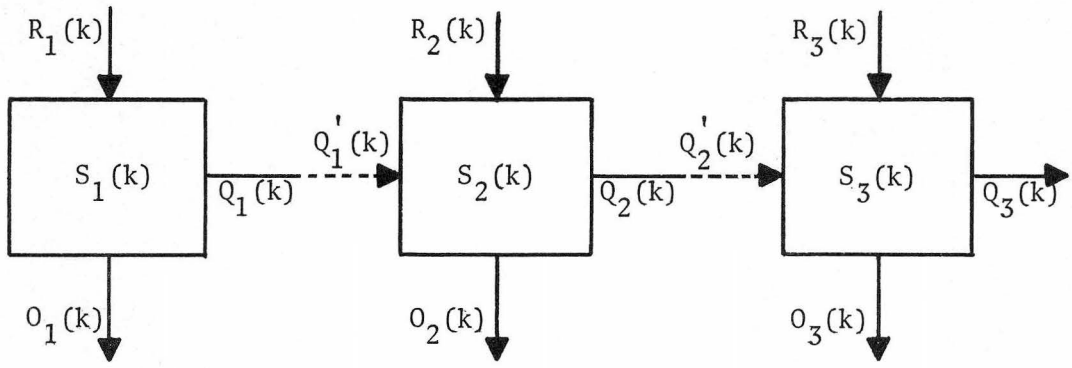

FIGURE IV-1
EXAMPLE THREE-RESERVOIR PROBLEM

$O_i(k)$ = average rate of overflow from reservoir i, during period k

$R_i(k)$ = average rate during period k of lumped direct stormflow input which is translated from the near vicinity of reservoir i [Note: assume that all direct input can be lumped, as shown in Figure IV-1, with negligible direct input occuring between reservoirs]

$Q_i(k)$ = average rate of throughput during period k, from reservoir i, with $Q_3(k)$ going to treatment [$i = 1,2$]

$Q_i'(k)$ = the routed or translated throughput from reservoir i, entering reservoir $i+1$ [$i = 1,2$]

$S_i(k)$ = storage in reservoir i, at the beginning of period k.

A common method of flow routing is the Muskingum method [12], where

$$Q_i'(k+1) = Q_i'(k) + T_i(Q_i'(k), Q_i(k), Q_i(k+1)) \qquad (2)$$
$$(k = 1,\ldots,M-1)$$

The transformation T_i may be linear or nonlinear, depending upon whether or not the coefficients associated with the Muskingum method are considered to be functions of flow rate. Backwater effects are not properly considered here, as in more realistic methods [12], but (2) will suffice for now.

We would like to formulate an optimization problem for minimizing $O(k)$, for all k, which is consistent with the general format of Problem B. This is hindered by the present form of Equation (2), since it is not consistent with the general format for dynamic or state equations (i.e., $k+1$ appears on the right-hand side). This can be remedied by defining a new state variable $V_i(k)$, and replacing (2) with [assuming $\Delta t = 1$]:

$$Q_i'(k+1) = Q_i'(k) + T_i(Q_i'(k), Q_i(k), V_i(k)) \tag{3}$$

$$Q_i(k+1) = Q_i(k) + [V_i(k) - Q_i(k)] \tag{4}$$
$$(k = 1,\ldots,M-1)$$

We see that (2) <=> (3),(4); $Q_i(k)$ is now regarded as a state variable, and $V_i(k)$ as a control variable, since $Q_i(k+1)$ is dependent upon $V_i(k)$.

We can now formulate the optimization problem [Problem D]:

$$\min_{\substack{S,Q,Q',\\ V,O}} \sum_{k=1}^{M} \omega_k \cdot O(k) \tag{5}$$

[where $S,Q \in E^{3(M+1)}, V \in E^{3M}, O \in E^{3M}$ and $Q' \in E^{2M}$]

subject to:

dynamic equations

$$\begin{cases} S_1(k+1) = S_1(k) + R_1(k) - Q_1(k) - O_1(k) & (6) \\ S_2(k+1) = S_2(k) + R_2(k) + Q_1'(k) - Q_2(k) - O_2(k) & (7) \\ S_3(k+1) = S_3(k) + R_3(k) + Q_2'(k) - Q_3(k) - O_3(k) & (8) \\ Q_i'(k+1) = Q_i'(k) + T_i(Q_i'(k), Q_i(k), V_i(k)) & (9) \\ Q_i(k+1) = Q_i(k) + [V_i(k) - Q_i(k)] & (10) \\ \qquad (i = 1,2 \qquad k = 1,\ldots,M) \end{cases}$$

initial conditions

$$\begin{cases} S_i(1) = S_i^o & \text{(given)} \\ Q_j'(1) = 0 & \text{(given)} \\ Q_i(1) = 0 \end{cases} \qquad i = 1,2,3;\ j = 1,2$$

state-space constraints

$$0 \leq S_i(k) \leq S_{i,max} \qquad i = 1,2,3 \qquad k = 2,\ldots,M \qquad (11)$$

$$0 \leq Q_j'(k) \leq Q_{j,max} \qquad j = 1,2 \qquad (12)$$

$$0 \leq Q_i(k) \leq Q_{i,max} \qquad (13)$$

final conditions

$$S_i(M+1) = 0 \qquad (14)$$

$$Q_j'(M+1) = 0 \qquad i = 1,2,3;\ j = 1,2 \qquad (15)$$

$$Q_i(M+1) = 0 \qquad (16)$$

control constraints

$$0 \leq V_i(k) \leq Q_{i,max} \qquad i = 1,2,3 \qquad (17)$$

$$O(k) \geq 0 \qquad k = 1,\ldots,M \qquad (18)$$

Physically speaking, Q_i and Q_i' are only defined for $k = 1,\ldots,M$. In order to be consistent with the general format, we are defining $Q_i(M+1) = Q_i'(M+1) \triangleq 0$, and expressing them as final conditions. Notice that (14) is arbitrary, and we could simply specify

$$S_i(M+1) \geq 0 \qquad (19)$$

as a more realistic final condition. This is, however, not consistent with the general format of Problem B. The transversality conditions for this situation are the same, however, with the exception that the multiplier ρ is now nonnegative.

$$u(k) \triangleq (V(k),O(k)) \qquad \Rightarrow m = 6$$

$$x(k) \triangleq (S(k),Q(k),Q'(k)) \qquad \Rightarrow n = 9$$

$$f(\cdot,u(k)) \triangleq \omega_k \cdot O(k)$$

$$\phi(x(M+1)) \triangleq 0$$

For this problem, then, there are a total of 15M variables (state and control) and 22M constraints, not including nonnegativity restrictions. Suppose, for example, that a storm lasts for about an hour, and control is exercised every 5 minutes. Then M = 12, the number of variables is 180, and there are 264 constraints. Therefore, optimization techniques applied to Problem D must involve some kind of decomposition strategy, where the original problem is replaced by several smaller problems.

C. DIRECT METHODS

C.1 Linear Programming

If the transformation T_i is linear, then Problem D can be solved by linear programming, and a global solution is assured if the problem is well-posed. It should be pointed out that in applying linear programming, there is no need to transform the problem into the format of Problem B (resulting in Problem D). The original routing relation (Equation (2)) suffices, and there is no need to add variables $V_i(k)$ and the constraints (4). The number of variables is therefore reduced to 12M, and the number of constraints to 16M. We will refer to this modified problem as Problem $\overline{D}$.

There are several ways of decomposing the linear programming problem, including Dantzig-Wolfe decomposition [21], generalized linear programming [14], and Rosen's partition programming [21]. Specialized techniques have been developed specifically for linear control problems [14], and have the advantage of being computationally efficient even if M is very large. Another technique, using generalized duality theory, that is discussed in a subsequent section, maybe applicable to linear control.

Considerable reduction in computational load can be gained by solving Problem $\overline{D}$ via its dual. For the primal problem, we must add [12M - 9] slack variables to the state-space and control constraints. Therefore,

a 16M × 16M basis must be used. For the dual problem, the basis will be 12M × 12M. Some computational results are reported in [30].

With respect to the questions asked in the introduction to this chapter, the linear programming approach can be evaluated as follows:

1. The linearity of T_i implies that flow routing is independent of flow rates in the sewer, which is not consistent with reality. The question is, what magnitude of error is introduced? If the error is tolerable, then a linear programming approach may be feasible. This error can only be evaluated, for particular problems, by comparing the linear routing with more realistic routing procedures.
2. As mentioned above, a number of efficient procedures are available for decomposing Problem $\overline{D}$ when M is large.
3. The linear programming algorithm assures convergence to a global solution, as long as the problem is well-posed.

C.2 Dynamic Programming

In applying dynamic programming to Problem D, we need not make any assumptions regarding T_i. The basic recursion relations are:

$$F_k(S(k),Q(k),Q'(k)) \tag{20}$$

$$= \min_{\substack{V(k),O(k),\\ [\text{subject to}\\ (6)-(10),\\ (11)-(13),\\ \text{and } (17),(18)]}} [\omega_k \cdot O(k) + F_{k+1}(S(k+1),Q(k+1),Q'(k+1))]$$

where (20) is applicable for $k = 1,\ldots,M-1$ (initial conditions for $k=1$ are given) and is solved for all combinations of $S(k),Q(k),Q'(k)$, with each variable taking on a finite number of values in the intervals $[0,S_{max}]$, $[0,Q_{max}]$, and $[0,Q_{max}]$, respectively. For $k = M+1$

$$F_M(S(M),Q(M),Q'(M)) \tag{21}$$

$$= \min_{V(M),O(M)} [\omega_M \cdot O(M)]$$

$$[\text{subject to (14) - (18)}]$$

It is immediately evident that Problem D cannot be solved by conventional dynamic programming, due to the so-called [8a] "curse-of-dimensionality." There are a total of 9 state variables at each stage. If the above intervals are discretized into L-1 subintervals, then a total of L^9 values must be stored at each stage, far exceeding the capacity of any computer, assuming L is a reasonable number. By "reasonable," we mean that it is large enough to insure a global solution.

The only way to obtain any kind of solution to (20) and (21) is to reduce L, which leads us into the possibility of applying incremental dynamic programming [20]. Here, $L = 3$, and we start with an initial guess for the optimal control $V^0(k),O^0(k))$ and perturb around the state variables resulting from these controls $(S^0(k),O^0(k),Q'^0(k))$. The perturbation is generally limited to one step above or below $S^0(k),Q^0(k),Q'^0(k)$, for all k, which explains why $L = 3$. A new optimal control $V^1(k),O^1(k)$ is then determined for this limited state-space, and the procedure is repeated until convergence is attained. Since we are making no assumptions regarding T_i, incremental dynamic programming can, in general, only insure convergence to local solutions.

Our evaluation of the applicability of dynamic programming is:

1. Realistic flow routing procedures may be utilized, since dynamic programming requires no restrictive assumptions regarding the transformation T_i.
2. Dynamic programming is a decomposition technique in itself, since the original problem is decomposed into a set of subproblems which are defined for each stage. The subproblems, however, become unwieldy from a computer storage standpoint (rapid-access storage) as the number of state variables increases. For only one reservoir, standard dynamic programming could be applied. For more than one, it must be abandoned.
3. For general problems involving more than one reservoir, incremental dynamic programming can be applied as a means of obtaining at least a local solution.

C.3 Nonlinear Programming Methods

All nonlinear programming codes require that assumptions (i) - (iii), in Section B of Chapter II, be applicable, in order for convergence to a global solution to be assured.† The objective function for our problem is, of course, linear, but a nonlinear transformation function T_i immediately violates assumption (iii). Since (3) is an equality constraint, the constraint region defined by (3) is convex if and only if (3) is linear [18]. Therefore, nonlinear programming methods are poorly suited for Problem D if a high value is placed on obtaining global solutions. Optimization problems with nonlinear equality constraints are generally considered to be the most difficult [18]. Our evaluation, then, is

---

†Note: that is, with the exception of direct enumeration or grid-search methods.

1. Most nonlinear programming codes require that T_i be at least continuous, in order for a solution to be obtained. Usually, stronger assumptions of differentiability are required for the algorithms to operate properly. These assumptions do not seem restrictive for utilization of realistic flow routing procedures.
2. The nonlinearity of T_i, and hence the nonconvexity of Problem D, limits the possible decomposition methods that could be applied. The two most important methods would probably be (i) Geoffrion's resource directive approach [16] and (ii) application of generalized duality theory [see Appendix].
3. In general, all standard nonlinear programming codes, that are not based on grid search methods operating over the entire constraint region of a problem, can at most guarantee convergence to local solutions.

D. AN INDIRECT METHOD - *THE APPROXIMATE-FLOW TECHNIQUE*

D.1 Introduction

In addressing ourselves to the most general combined sewer problem for a particular subbasin, where

(i) there are several interconnected storage basins

(ii) realistic routing procedures are utilized, thus introducing nonlinearities into the state equations, and therefore resulting in nonconvexity of the control problem,

we can conclude that direct application of

1. linear programming is not possible, unless some kind of linearization procedure is carried out. In general, though,

global solution of the original nonlinear problem is difficult to attain by these methods

2. standard dynamic programming, though being a global solution technique, is not feasible, due to dimensionality difficulties caused by interconnection of several reservoirs. Incremental dynamic programming is applicable, but can only guarantee local solutions.
3. nonlinear programming methods can only assure convergence to local solutions.

The following technique may be an answer to both the problem of obtaining global solutions and the dimensionality difficulties encountered in applying dynamic programming.

D.2 Approximation of Routed Flow

Suppose we are given arbitrary functions $\phi(\alpha,t)$, $\psi(\beta,t)$, with parameters α,β, respectively, such that if $Q'^{*}(k)$, $k = 1,\ldots,M$, is the global solution (assuming that it is unique) to Problem D, then there exist α^{*},β^{*} such that

$$\begin{aligned} Q_1'^{*}(k) &= \phi(\alpha^{*},t_k) \\ Q_2'^{*}(k) &= \psi(\beta^{*},t_k) \end{aligned} \qquad k = 1,\ldots,M$$

Then, the following problem can be written which is exactly equivalent to Problem D [Problem E1]:

$$\min_{\substack{S,Q,Q',\\ V,O,\alpha,\beta}} \sum_{k=1}^{M} \omega_k \cdot O(k) \tag{22}$$

subject to:

state equations

$$\begin{cases} S_1(k+1) = S_1(k) + R_1(k) - Q_1(k) - O_1(k) & (23) \\ S_2(k+1) = S_2(k) + R_2(k) + \phi(\alpha,t_k) - Q_2(k) - O_2(k) & (24) \\ S_3(k+1) = S_3(k) + R_3(k) + \psi(\beta,t_k) - Q_3(k) - O_3(k) & (25) \\ Q_i'(k+1) = Q_i'(k) + T_i(Q_i'(k),Q_i(k),V_i(k)) & (26) \\ Q_i(k+1) = Q_i(k) + [V_i(k) - Q_i(k)] & (27) \\ \quad (i = 1,2; \quad k = 1,\ldots,M) \end{cases}$$

initial conditions

$$S(1),Q'(1),Q(1) \quad \text{(given)}$$

state-space constraints

$$\begin{cases} S(k) \leq S_{max} & (28) \\ Q'(k) \leq Q_{max} & (29) \\ \phi(\alpha,t_k) - Q_1'(k) = 0 & (30) \\ \psi(\beta,t_k) - Q_2'(k) = 0 & (31) \end{cases} \qquad k = 2,\ldots,M$$

final conditions

$$\begin{cases} S(M+1) = 0 \\ Q'(M+1) = 0 \\ Q(M+1) = 0 \end{cases}$$

control constraints

$$Q(k) \leq Q_{max}, \quad k = 1,\ldots,M \qquad (32)$$

where $S(\cdot),Q(\cdot),O(\cdot),S_{max},Q_{max},\omega_k \in (E^3)^+; Q'(\cdot) \in (E^2)^+; \alpha \in E^a, \beta \in E^b$.

Problem E1 is not in the strict format of Problem D, due to the inclusion of α,β as variables and the form of the state-space constraints. We will see, however, how Problem E1 can be decomposed into subproblems which are consistent with the general format. Notice, also, that each of the state-space constraints can be replaced by two inequalities. The addition of (30) and (31) insures the equivalence of D and E1, since we have assumed that an exact fit can be made between the $Q_i'^*(k)$ $(k = 1,\ldots,M)$ and given functions $\phi(\alpha^*,t_k)$, $\psi(\beta^*,t_k)$.

In general, however, it is impossible to find a function that will give an exact fit. Thus, solution of Problem E1 would produce solutions Q'^{*} that would be suboptimal, since (30) and (31) must be satisfied. The closeness of the fits would tend to increase as a and b increase. But since α and β are now variables, we desire to limit their sizes to a degree that will facilitate solution. Taking this into account, we would like to modify E1 in the following way [Problem E2]:

$$\min_{\substack{S,Q,Q',\\ V,O,\alpha,\beta}} \sum_{k=1}^{M} [\omega_k \cdot O(k) + \mu_k[\phi(\alpha,t_k) - Q_1'(k)]^2 + [\psi(\beta,t_k) - Q_2'(k)]^2] \tag{33}$$

subject to:

[all the constraints of Problem E1, with the exception of (30) and (31)]

There are no weighting factors attached to the 3rd term of the objective function, since proper adjustment of the ω_k, μ_k can produce any desired relative weighting among all the terms. For Problem E2, we can now allow some error in the fitting process. As long as this error is tolerable, then even though Problem E2 is no longer exactly equivalent to E1, for all practical purposes, we can replace E1 with E2.

The advantage of placing the problem in the form of E2 is that we can now decompose the problem into a set of dynamic programming problems, for which dimensionality is no longer a great problem. This is accomplished by the powerful *projection theorem* [16], which states that we can write Problem E2 in the following equivalent form [Problem E3]:

$$\min_{\alpha,\beta} \; v(\alpha,\beta) \tag{34}$$

where

$$v(\alpha,\beta) = \min_{\substack{S,Q,Q',\\ V,O}} \sum_{k=1}^{M} [\omega_k \cdot O(k) + \mu_k[\phi(\alpha,t_k) - Q_1'(k)]^2 + [\psi(\beta,t_k) - Q_2'(k)]^2] \tag{35}$$

subject to:

[the constraints associated with Problem E2]

The problem has now been *projected* onto the space of the α,β. The minimization carried out in (35) is based on given α,β, and is referred to as the *inner problem*. The *outer problem* is represented in (34). Strictly speaking, we should replace *min* in (35) with *inf* (infimum, or greatest lower bound), but (35) holds as long as we assume that a minimum exists in (35) for all $\alpha \in E^a$, $\beta \in E^b$. To assure this, it may be necessary to place arbitrary upper and lower bounds on α,β.

With α,β now representing given parameters for the inner problem, the inner problem is now completely decomposible into three, independent, three-dimensional dynamic programming subproblems. In general,

Subproblem i (for $i = 1,2,3$) is

$$\min_{\substack{S_i,Q_i,Q_i',\\ V_i,O_i}} \sum_{k=1}^{M} [\omega_{1k} O_1(k) + TERM_i] \tag{36}$$

where

$$\text{for } i = 1, \quad TERM_i = \mu_k[\phi(\alpha,t_k) - Q_1'(k)]^2$$

$$\text{for } i = 2, \quad TERM_i = [\psi(\beta,t_k) - Q_2'(k)]^2$$

$$\text{for } i = 3, \quad TERM_i = 0$$

subject to:

[(6) - (18), for appropriate i]

Each of these subproblems is solved by dynamic programming in the format of (20) and (21), except that the state and control vectors

depicted in (20) and (21) are replaced with their respective components for the ith subproblem.

In summary, we have replaced the original 9-dimensional dynamic programming problem (which was impossible to solve) with an outer, unconstrained nonlinear programming problem involving a total of $a + b$ variables, which at each iteration solves a total of three, three-dimensional dynamic programming subproblems. If the dynamic programming subproblems are efficiently programmed, then computer storage should no longer be a problem. There are, however, two difficulties that present themselves:

(a) Since the 3-dimensional dynamic programming subproblems are solved a number of times, depending on how long it takes the outer problem to converge, the amount of computer time required will render solution by this approach infeasible.

(b) As we begin to consider more complex subbasin configurations involving many reservoirs, the outer problem will involve a quantity of variables that increases in proportion to the number of reservoirs. Hence, it will become more and more difficult to insure convergence to global solution of the outer problem. For nonconvex problems, the only sure way of finding global solutions is via grid-search techniques. As the dimensionality of the outer problem increases, grid-search becomes less feasible, due to the enormous number of computations involved in any direct enumeration procedure.

The above discussion points to two goals:

(i) Somehow reduce the computational effort involved in the inner problem.

(ii) Provide some means of keeping the outer problem of reasonable size.

The following section addresses itself to (i), and opens the way to consideration of (ii).

D.3 Approximation of Flow prior to Routing

Suppose that instead of approximating the routed flows $Q_i'(\cdot)$ by the functions $\phi(\cdot,\cdot)$ and $\psi(\cdot,\cdot)$, for i = 1,2, respectively, we approximate the throughflows $Q_i(\cdot)$ prior to routing. Then the following problem can be written which is exactly equivalent to Problem D [Problem F1]:

$$\min_{\substack{S,Q',\\ Q,O,\alpha,\beta}} \sum_{k=1}^{M} \omega_k \cdot O(k) \tag{37}$$

subject to:

state equations

$$S_1(k+1) = S_1(k) + R_1(k) - Q_1(k) - O_1(k) \tag{38}$$

$$S_2(k+1) = S_2(k) + R_2(k) + Q_1'(k) - Q_2(k) - O_2(k) \tag{39}$$

$$S_3(k+1) = S_3(k) + R_3(k) + Q_2'(k) - Q_3(k) - O_3(k) \tag{40}$$

$$Q_1'(k+1) = Q_1'(k) + T_1(Q_1'(k), \phi(\alpha,t_k), \phi(\alpha,t_{k+1})) \tag{41a}$$

$$Q_2'(k+1) = Q_2'(k) + T_2(Q_2'(k), \psi(\beta,t_k), \psi(\beta,t_{k+1})) \tag{41b}$$

$$(k = 1,\ldots,M)$$

initial conditions

$$S(1), Q'(1) \quad \text{(given)}$$

state-space constraints

$$0 \leq S(k) \leq S_{max} \tag{42}$$

$$0 \leq Q'(k) \leq Q_{max} \qquad k = 2,\ldots,M \tag{43}$$

final conditions

$$S(M+1) = 0$$

$$Q'(M+1) = 0$$

$$\text{control constraints} \begin{cases} 0 \leq Q(k) \leq Q_{max} & (44) \\ \phi(\alpha,t_k) - Q_1(k) = 0 \qquad k = 1,\ldots,M & (45) \\ \psi(\beta,t_k) - Q_2(k) = 0 & (46) \end{cases}$$

As before, in considering that it is generally impossible to obtain an exact fit, as required in (45) and (46), we apply the projection theorem to a modification of Problem F1, where (45) and (46) are deleted as constraints and the objective function includes the fitting error as an optimality criterion [Problem F2]:

$$\min_{\alpha,\beta} \; v(\alpha,\beta)$$

where

$$v(\alpha,\beta) = \min_{\substack{S,Q',\\ Q,O}} \sum_{k=1}^{M} [\omega_k \cdot O(k) + \mu_k [\phi(\alpha,t_k) - Q_1(k)]^2 \; [\psi(\beta,t_k) - Q_2(k)]^2 \qquad (48)$$

subject to:

[the constraints associated with Problem F1, with (45) and (46) deleted]

Again, the inner problem (48) is decomposable into three independent subproblems, corresponding to each reservoir. The importance difference from the previous formulation (Problem E3) is that these subproblems can be solved as *one-dimensional* dynamic programming problems. In the previous formulation, we approximated the dependent or state variable Q', so that it was required to retain it in the subproblems. In this present formulation, we are approximating the independent or control variable Q, so that given α,β from the outer problem, Q' is uniquely specified as a function of α,β. Since Q' no longer appears in the objective function, we can delete it from our problem. This is accomplished by noting that (41) can be written in the following equivalent forms [assuming $Q_i'(1)$ given]:

$$Q_1'(2) = Q_1'(1) + T_1(Q_1'(1),\phi(\alpha,t_1),\phi(\alpha,t_2))$$
$$Q_1'(3) = Q_1'(1) + T_1(Q_1'(1),\phi(\alpha,t_1),\phi(\alpha,t_2)) + T_1((Q_1'(1) + T_1(Q_1'(1),\phi(\alpha,t_1),\phi(\alpha,t_2)), \phi(\alpha,t_2),\phi(\alpha,t_3))$$
$$\vdots$$

(49a)

$$Q_1'(M+1) = Q_1'(1) + \sum_{k'=1}^{M} T_1(Q_1'(k'),\phi(\alpha,t_{k'}),\phi(\alpha,t_{k'+1})) + T_1((Q_1'(1) + \sum_{k'=1}^{M} T_1(Q_1'(1),\phi(\alpha,t_{k'}),\phi(\alpha,t_{k'+1})), \phi(\alpha,t_M),\phi(\alpha,t_{M+1}))$$

[the relations for $Q_2'(k)$, $k = 1,\ldots,M$ are of the same form, except $\phi(\alpha,t_k)$ is replaced with $\psi(\beta,t_k)$, for all k] (49b)

Since the vector $Q'(k)$ has now been represented as a function of α,β (referred to as $Q'(\alpha,\beta,k)$), let us modify F2 in such a way that $Q'(\alpha,\beta,k)$ is deleted from the inner problem [Problem F3]:

$$\min_{\alpha,\beta} \; v(\alpha,\beta,Q'(\alpha,\beta,\cdot)) \tag{50}$$

subject to:

(49a), (49b), and (43)

where $Q'(\alpha,\beta,k) = (Q_1'(\alpha,k),Q_2'(\beta,k))$, and

$$v(\alpha,\beta,Q'(\alpha,\beta,\cdot)) = \min_{S,Q,O} \sum_{k=1}^{M} [\omega_k \cdot O(k) + \mu_k[\phi(\alpha,t_k) - Q_1(k)]^2 + [\psi(\beta,t_k) - Q_2(k)]^2] \tag{51}$$

subject to:

[the constraints associated with Problem F1, with (43), (45), and (46) deleted]

The inner problem (51) can now be decomposed into the following subproblems:

Subproblem 1:

$$\min_{S_1,Q_1,O_1} \sum_{k=1}^{M} [\omega_{1k} O_1(k) + \mu_k [\phi(\alpha,t_k) - Q_1(k)]^2$$

subject to: (all variables assumed nonnegative)

$$S_1(k+1) = S_1(k) + R_1(k) - Q_1(k) - O_1(k)$$

$$S_1(1), S_1(M+1) \quad \text{(given)}$$

$$S_1(k) \leq S_{1,\max}, \quad k = 2,\ldots,M$$

$$Q_1(k) \leq Q_{1,\max}, \quad k = 1,\ldots,M \bullet$$

which is easily solved as a one-dimensional dynamic programming problem ($S_1(k)$ as the state variable) with two decision variables at each stage k ($Q_1(k), O_1(k)$).

Subproblem 2:

$$\min_{S_2,Q_2,O_2} \sum_{k=1}^{M} [\omega_{2k} O_2(k) + [\psi(\beta,t_k) - Q_2(k)]^2]$$

subject to:

$$S_2(k+1) = S_2(k) + R_2(k) + Q_1'(\alpha,k) - Q_2(k) - O_2(k)$$

$$S_2(1), S_2(M+1) \quad \text{(given)}$$

$$S_2(k) \leq S_{2,\max}, \quad k = 2,\ldots,M$$

$$Q_2(k) \leq Q_{2,\max}, \quad k = 1,\ldots,M$$

where Q' have been given as a parameters from the outer problem.

Subproblem 3:

$$\min_{S_3,Q_3,O_3} \sum_{k=1}^{M} [\omega_{3k} O_3(k)]$$

subject to:

$$S_3(k+1) = S_3(k) + R_3(k) + Q_2'(\beta,k) - Q_3(k) - O_3(k)$$

$$S_3(1), S_3(M+1) \quad \text{(given)}$$

$$S_3(k) \leq S_{3,max}, \quad k = 2,\ldots,M$$

$$Q_3(k) \leq Q_{3,max}, \quad k = 1,\ldots,M$$

All of the above subproblems are solveable as one-dimensional dynamic programming problems, with two control or decision variables at each stage. We can further reduce the computation time by utilizing the formulation discussed in Section A.1, Chapter II. That is, we let

$$\overline{Q}(k) = Q(k) + O(k)$$

and replace the overflow terms in the objective functions for the subproblems with

$$O(k) = [S(k) - S_{max}] \tag{52}$$

for all $k \in K_i$, where

$$K_i = \{k | S_i(k) - S_{i,max} \geq 0\}, \quad i = 1, 2, 3 \tag{53}$$

so that we have only one control variable $\overline{Q}_i(k)$ for subproblem i, at each stage k, and the upper bound S_{max} is ignored.

We have, therefore, realized goal (i), given in Section D.2. Our attention now focuses on goal (ii). Before discussing ways of meeting this goal, it should be noted that the above method of reducing the number

of control variables at each stage from two to one, though resulting in less computation time for the inner problem, introduces additional nonconvexity into the outer problem, thus making it more difficult to find a global solution. The proper trade-off between time and nonconvexity can only be resolved through extensive computational experience.

D.4 Application of Generalized Duality Theory

Having found ways of significantly lessening the computational burden associated with the inner problem, we now address ourselves to global solution of the outer problem. We see from Problem F3 that as the number of reservoirs increases to $N > 3$, then the number of variables associated with the outer problem increases to approximately $\bar{a} \times N$, where $\bar{a}$ represents the average number of components of the parameter vectors associated with the functions approximating throughflow Q.

For illustrative purposes, let us return to Problem F3 (where $N = 3$), and place it in the following equivalent form [Problem G]:

$$\min_{\alpha,\beta,\alpha',\beta'} \quad v(\alpha,\beta,\alpha',\beta',Q'(\alpha',\beta',k)) \tag{54}$$

subject to:

(49) and (43) (with α,β replaced by α',β', respectively), plus

$$\alpha - \alpha' = 0 \tag{55}$$

$$\beta - \beta' = 0 \tag{56}$$

where $\alpha,\alpha' \in E^a; \beta,\beta' \in E^b$, and

$$v(\alpha,\beta,\alpha',\beta',Q'(\alpha',\beta'))$$

$$= \min_{S,Q,O} \sum_{k=1}^{M} [\omega_k \cdot O(k) + \mu_k[\phi(\alpha,t_k) - Q_1(k)]^2 + [\psi(\beta,t_k) - Q_2(k)]^2] \tag{57}$$

subject to:

[the constraints of F1 with (43), (45), and (46) deleted]

Notice that Q' has been expressed as a function of α',β', but that addition of (55) and (56) preserves the equivalence of F3 and G. The advantage of using Problem G over Problem F3 is that we can place (55) and (56) into (54) via introduction of *generalized Lagrange multipliers* λ, as discussed at length in the Appendix. Hence, we write the *dual function* (see Appendix) associated with the outer problem of Problem G or

$$h(\lambda) = \min_{\alpha,\beta,\alpha',\beta'} \{v(\alpha,\beta,\alpha',\beta',Q'(\alpha',\beta',k)) + \lambda_1 \cdot [\alpha - \alpha'] + \lambda_2 \cdot [\beta - \beta']\} \tag{58}$$

where $\lambda \in E^a \times E^b$, or $\lambda \in E^{a+b}$

subject to:

(49) and (43)

which is decomposable as follows

$$h(\lambda) = \{\min_{\alpha} [v_1(\alpha) + \lambda_1 \cdot \alpha] + \min_{\alpha',\beta} [v_2(\alpha',\beta) - \lambda_1 \cdot \alpha' + \lambda_2 \cdot \beta] + \min_{\beta'} [v_3(\beta') - \lambda_2 \cdot \beta]\} \tag{59}$$

where

$v_1(\alpha)$ is the solution of Subproblem 1 (stated in Section D.3)

$v_2(\alpha', \beta)$ is the solution of Subproblem 2, with α' replacing α.

$v_3(\beta')$ is the solution of Subproblem 3, with β' replacing β.

It is seen in (59) that for given λ, the outer problem can be expressed as three independent *sub-outer problems*, each involving no more than $a + b$ variables. In extending this method to N reservoirs, there will in turn be N sub-outer problems, each involving approximately $2\bar{a}$ variables where $\bar{a}$ represents the average number of parameters associated with each approximating function utilized.

The λ's must now be properly adjusted until (55) and (56) are satisfied which is hopefully accomplished by solving the *dual problem*. Such will be the case if a *saddle-point* exists [see Appendix]. Solution of the dual problem then indirectly solves the original N reservoir problem (by Theorems 1 and 2 in the Appendix). The dual problem, for the example three-reservoir problem, is

$$\max_{\lambda \in E^{a+b}} \quad h(\lambda) \tag{60}$$

and for the general N reservoir problem

$$\max_{\lambda \in E^{\bar{a}N}} \quad h(\lambda) \tag{61}$$

where $\bar{a}N$ is *approximately* the dimension of λ.

On the surface, it seems that we have accomplished little, in that even though the outer problem has been decomposed, it in turn has been imbedded in a dual problem which involves a large number of dual variables ($\cong \bar{a}N$). The advantage is that the dual problem is guaranteed to always be concave, no matter how nonconvex the sub-outer problems happen to be (by theorem 6 in the Appendix). By keeping the dimensionality of these sub-outer problems V_i to a restricted level, we increase the assurance of finding global solutions for them.

D.5 Discussion

Let us summarize the conditions presented in the Appendix which give assurance that solution of the dual problem (61) can be found, and that this solution will indirectly solve the original N-reservoir problem, as represented in Problem E1 for $N = 3$:

1. The vectors $\alpha,\beta,\alpha',\beta',S,Q,Q'$, and O must be contained in closed and bounded sets.
2. The objective functions associated with the subproblems given in Section D.3 must be continuous functions of all of the above variables.
3. For all given λ in the dual problem, the dual function must yield *g-unique* (which is a generalization of uniqueness) solutions $\alpha^*,\beta^*,\alpha'^*,\beta'^*,S^*,Q^*,Q'^*$, and O^*.
4. There exists a finite λ^* such that

$$\frac{\partial h(\lambda^*)}{\partial \lambda_i} = 0$$

for $i = 1,\ldots,\bar{a}N$.

Condition 1 is obviously satisfied for our problem, as long as we place an arbitrary upper bound on O, and arbitrary upper and lower bounds

on $\alpha,\beta,\alpha',\beta'$, such that the optimal solutions are contained in the interiors of these intervals. Condition 2 is satisfied as long as the functions approximating Q ($\phi(\cdot,\cdot),\psi(\cdot,\cdot)$, $\cdots$, etc.) are continuous functions of their respective parameters. Conditions 3 and 4 are more difficult to assure. It should be pointed out, however, that these are only *sufficient conditions*, and that a saddle-point may exist even though they are not strictly satisfied.

As stated previously, the goal of the dual problem is to adjust the λ until (55) and (56) are satisfied indirectly (this corresponds to maximizing the dual function $h(\lambda)$). If conditions 1 and 2 above are not strictly satisfied, then there may exist no λ^* such that (55) and (56) are exactly satisfied. This corresponds to a *duality gap*, and implies that a *saddle-point* does not exist [see Appendix]. It may be, however, that there exists a λ^* such that (55) and (56) are *almost* satisfied, or

$$|\alpha^*(\lambda^*) - \alpha'^*(\lambda^*)| = \varepsilon$$

$$|\beta^*(\lambda^*) - \beta'^*(\lambda^*)| = \sigma$$

where vectors ε and σ have tolerably small components. In this case, the duality gap is considered negligible, and a saddle-point is assumed to exist. These questions cannot be fully resolved without extensive computational experience.

A suggested algorithm for solving the dual problem (61) follows:

(a) Adjust the λ's for the dual problem utilizing a rapidly converging unconstrained maximization algorithm (e.g., Davidon-Fletcher-Powell, Powell's method, or steepest ascent [18]). If the method requires derivatives, assume that the dual function

is differentiable, and use the gradients defined in the Appendix.

(b) Ideally, for each given λ in the dual problem, global solutions to the sub-outer problems should be found. If $\bar{a} \leq 3$, then grid-search methods can probably be used. For $\bar{a} > 3$, constrained minimization methods can be used until the dual problem begins to converge. At this point, greater attempts should be made at attaining global solutions.

(c) The subproblems associated with the inner problem are easily solved via one-dimensional dynamic programming. Since they will be solved numerous times as λ, and in turn, $\alpha,\beta,\alpha',\beta'$ are adjusted, it is extremely important that computer codes be written as efficiently as possible.

In addition to the difficulty of assuring that Conditions 3 and 4 above are satisfied, there is the problem of finding approximation functions $\phi(\cdot,\cdot),\psi(\cdot,\cdot), \cdots$, etc., which will give accurate fits, while utilizing as few parameters as possible, so that global solution of the outer problem is more easily attained.

Aside from these difficulties and uncertainties in applying the dual approach, using flow approximation, the following advantages are clear:

1. There is potential for being able to obtain a solution which is assured to be the global solution. Such assurance is generally never possible when directly applying nonlinear programming algorithms to nonconvex problems such as this.

2. Even if a saddle-point does not exist, or there exists no λ^* such that (55) and (56) are satisfied, the amount of infeasibility may be negligible, for practical purposes.
3. For larger duality gaps, the infeasible solutions may be useful for generating accurate initial approximations for initiating a direct nonlinear programming code.

V. SUMMARY AND CONCLUSIONS

The optimal control problem associated with automated operation of ambient and/or auxiliary storage capabilities within combined sewer systems can be formulated as either a finite-dimensional (discrete-time) or infinite-dimensional (continuous-time) optimization problem. Both involve discretization at some stage, since digital computers can only deal with finite quantities of real numbers. For the former, discretization is carried out prior to problem solution, whereas for the latter it is effected during and subsequent to computation, since actual control of the system is carried out in discrete-time.

It was concluded that finite-dimensional optimization (FDO) is preferable to infinite-dimensional optimization (IDO) for the combined sewer problem, due to the following factors:

1. Actual operation of the system is carried out in discrete real-time. The size of FDO problems can be unwieldy if the time intervals are too small, so that IDO may be the only alternative. It appears, however, that intervals will be of moderate size, due mainly to the need for collecting and analyzing adequate quantities of sensor data in these intervals, for reasonable storm and flow prediction.
2. IDO is based on solving necessary conditions for optimality, which apply at solutions other than the desired global solution. FDO relies less on necessary conditions.
3. In general, for nonlinear problems, it is easier to obtain at least local solutions by FDO than IDO. It was shown that the necessary conditions for IDO can be derived as limiting cases of the necessary conditions for FDO. But there are difficulties in solving the former that do not arise in the latter.

4. In applying IDO, a continuous curve must be fitted to discrete rainfall data. Since there are an infinite number of such curves, the question of uniqueness of solutions arises.

These conclusions seem to be supported by computational experience. Applications of IDO to ambient storage models failed to give solutions in most cases, even though the flow model and system configuration was extremely idealized. This can be contrasted with the ease of obtaining results by linear programming for a comparably simple flow model and auxiliary storage configuration, as reported in [30]. There is some question, however, about the validity of comparing these results, since the ambient storage model required solution of more complicated equations, even though the flow routing assumptions were of comparable simplicity. As discussed in Chapter III, however, it seems possible to treat the ambient case from an auxiliary storage viewpoint, though no computational results are available as yet.

Turning to FDO, it was shown that linear flow routing models (e.g., the Muskingum method with constant coefficients) resulted in a large-scale linear programming problem, for which there are a number of efficient decomposition strategies available. If the error introduced by linear routing is tolerable, linear programming may be feasible for effective on-line optimization, since global solutions to linear problems are assured (under mild assumptions) in a finite number of iterations, by the simplex method.

Introduction of any degree of nonlinearity in the flow routing method (e.g., the Muskingum method with variable coefficients) results in a nonconvex FDO problem. Dynamic programming can deal with the nonconvexity problem, but the so-called curse-of-dimensionality precludes its applicability.

Incremental dynamic programming is a possibility, but can only give local solutions, in general. Nonlinear programming algorithms also suffer from the fact that convergence is generally to local solutions. Even if a global solution happens to be determined, there is no known way of verifying its globality, other than by inefficient direct enumeration.

In order to deal particularly with the problem of finding global solutions, an approximate-flow technique was developed which, in conjunction with generalized duality theory and the projection theorem, resulted in one-dimensional dynamic programming problems imbedded in constrained nonlinear programming problems of limited dimension, which in turn were imbedded in a dual problem for which global solution is assured as long as global solutions can be obtained for the interior subproblems. The dual problem solves (globally) the original control problem if and only if a saddle-point exists. If a saddle-point does not exist (which is not determinable *a priori* for nonconvex problems), an infeasible solution to the control problem results. If the infeasibility is of tolerable magnitude, then this solution will be adequate. Otherwise, the infeasible solution may be used to generate accurate initial approximations for direct application of constrained nonlinear programming algorithms.

Considerable computational experience is necessary in order to verify the applicability of the approximate-flow technique. It appears, though, that this method opens the way for finding global solutions to the nonconvex control problems resulting from realistic flow routing procedures. The goal is to obtain considerable off-line optimization results based on a large variety of historical and synthetically generated storm situations, so that optimal rule curves and operating policies can be programmed into

the on-line computer system. These policies can perhaps be utilized in conjunction with on-line optimization by linear programming. Though simplified linear flow models are required for the latter, on-line optimization has the advantage of being able to respond to the uniqueness of the particular storm event occurring in real time, which is not possible if all optimization is carried out off-line.

REFERENCES

[1] _____, Combined Sewer Overflow Abatement Technology, Water Pollution Control Series 11024, Federal Water Quality Administration, June 1970.

[2] Anderson, J. J. and D. J. Anderson, "Computer Control and Modeling of Sewer Systems," 73rd National Meeting, AICE, August 27-30, 1972.

[3] Banerjee, K., "Generalized Lagrange Multipliers in Dynamic Programming," Operations Research Center Report # ORC 71-12, University of California, Berkeley, June 1971.

[4] Bell, W., "Progress Report on Work Since Completion of Task IV Report," Department of Civil Engineering, Colorado State University, December 1972.

[5] Bell, W. and B. Wynn, "Minimization of Pollution from Combined Sewer Systems," paper presented at the International Symposium on Systems Engineering and Analysis, Purdue University, October 1972.

[6] Bell, W., G. Johnson, and B. Wynn, "Simulation and Control of Flow in Combined Sewers," Sixth Annual Simulation Symposium, Tampa, Florida, March 1973.

[7] Bell, W., B. Wynn, and G. L. Smith, "Model of Real-Time Automation and Control Systems," Technical Report #7, OWRR - Metropolitan Water Intelligence Systems Project, Department of Civil Engineering, Colorado State University, February 1972.

[8] Bellman, R. E. and R. E. Kalaba, QUASILINEARIZATION AND NONLINEAR BOUNDARY-VALUE PROBLEMS, American Elsevier, 1965.

[8a] Bellman, R. E. and S. E. Dreyfus, APPLIED DYNAMIC PROGRAMMING, Princeton University Press, 1962.

[9] Bryson, A. E. and Y. C. Ho, APPLIED OPTIMAL CONTROL, Blaisdell, 1969.

[10] Canon, M. D., C. D. Callum, and E. Polak, THEORY OF OPTIMAL CONTROL AND MATHEMATICAL PROGRAMMING, McGraw-Hill, 1970.

[11] Citron, Stpehen J., ELEMENTS OF OPTIMAL CONTROL, Holt, Rinehart and Winston, 1969.

[12] Cunge, J. A., "Mathematical Modeling of Open Channel Flow," unpublished lecture notes, Department of Civil Engineering, Colorado State University, 1973.

[13] Dantzig, G. B., "Linear Control Processes and Mathematical Programming," SIAM Journal of Control, Vol. 4, 1966.

[14] Dantzig, G. B. and R. M. Van Slyke, "Generalized Linear Programming," in OPTIMIZATION FOR LARGE-SCALE SYSTEMS... WITH APPLICATIONS, D. A. Wismer (ed.), McGraw-Hill, 1971.

[15] Field, Richard, "Management and Control of Combined Sewer Overflows: Program Overview," paper presented at the 44th Annual Meeting of the New York Water Pollution Control Association, January 26-28, 1972.

[16] Geoffrion, A. M., "Large-Scale Linear and Nonlinear Programming," in OPTIMIZATION METHODS FOR LARGE-SCALE SYSTEMS... WITH APPLICATIONS, D. A. Wismer (ed.), 1971.

[17] Gibbs, C. V., S. M. Stuart, and C. B. Curtis, "System for Regulation of Combined Sewage Flows," Journal of the Sanitary Division, ASCE, December 1972.

[18] Himmelblau, D. M., APPLIED NONLINEAR PROGRAMMING, McGraw-Hill, 1972.

[19] Jizmagian, S., "Generalized Programming Solution of Continous-Time Linear System Optimal Control Problems," doctoral dissertation, Operations Research Department, Stanford University, 1968.

[19a] Karlin, S., MATHEMATICAL METHODS AND THEORY IN GAMES, PROGRAMMING, AND ECONOMICS, Vol. 1, Addison-Wesley, 1959.

[19b] Labadie, J. W., "Decomposition of a Large-Scale, Nonconvex Parameter Identification Problem in Geohydrology," Operations Research Center Report # ORC 72-12, University of California, Berkeley, September 1972.

[20] Larson, R. E., STATE INCREMENT DYNAMIC PROGRAMMING, American Elsevier, 1968.

[21] Lasdon, Leon S., OPTIMIZATION THEORY FOR LARGE SYSTEMS, MacMillan, 1970.

[21a] Luenberger, David G., OPTIMIZATION BY VECTOR SPACE METHODS, Wiley, 1969.

[22] McPherson, Murray B., "Feasibility of the Metropolitan Water Intelligence System Concept," Technical Memorandum #15, ASCE Urban Water Resources Research Program, December 1971.

[23] Mesarovic, M. D., D. Macko, and Y. Takahara, THEORY OF HIERARCHICAL, MULTILEVEL, SYSTEMS, Academic Press, 1970.

[24] Noton, Maxwell, INTRODUCTION TO VARIATIONAL METHODS IN CONTROL ENGINEERING, Pergamum Press, 1965.

[25] Noton, Maxwell, MODERN CONTROL ENGINEERING, Pergamum Press, 1972.

[25a] Rockafeller, R. T., CONVEX ANALYSIS, Princeton University Press, 1970.

[26] San Francisco Department of Public Works, "San Francisco Master Plan for Wastewater Management," 1971.

[27] Varaiya, P. P., NOTES ON OPTIMIZATION, Van Nostrand Reinhold, 1972.

[28] Wilde, D. J. and C. S. Beightler, FOUNDATIONS OF OPTIMIZATION, Prentice-Hall, 1967.

[29] Zangwill, Willard I., NONLINEAR PROGRAMMING: A UNIFIED APPROACH, Prentice-Hall, 1969.

[30] Ninth Quarterly Report to the Office of Water Resources Research, Metropolitan Water Intelligence Systems Project; M. L. Albertson, N. S. Grigg, and G. L. Smith, Co-Principal Investigators; Department of Civil Engineering, Colorado State University, January 1973.

APPENDIX

SUMMARY OF GENERALIZED DUALITY THEORY

The following is a concise review of the basic concepts and results of generalized duality theory, and is taken from [19b]. As discussed in Section D.4, Chapter IV, application of generalized duality theory opens the way to dealing with the complex, large-scale nature of the combined sewer problem, which arises even in subbasin analysis. In particular, there is potential for indirectly finding global solutions to the large-scale nonconvex control problem discussed in Chapter IV, whereas direct nonlinear programming techniques can generally only find local solutions. The great advantage of the dual approach is that its solution will either give the global solution desired, or give an infeasible solution, under certain mild assumptions. If the infeasibility is of small degree, then this solution will suffice. Direct methods, on the other hand, produce solutions which are generally impossible to define as being local or global.

Most of this material is condensed from excellent presentations by Lasdon [21], Banerjee [3], and Varaiya [27]:

Given the *primal problem*

$$\min_{x \in X} \; f(x)$$

subject to

$$g_i(x) \leq 0$$

$$(i = 1,\ldots,m)$$

where $x = (x_1,\ldots,x_n)$; X is a subset of E^n, we can write the *Lagrangian function* as

$$L(x,\lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x), \quad \text{for} \quad \lambda_i \geq 0$$

A point (x^o,λ^o), $\lambda^o \geq 0$, $x^o \epsilon X$, is a *saddle-point* for L if it satisfies

(i) $L(x^o,\lambda^o) \leq L(x,\lambda^o)$, for all $x \epsilon X$

(ii) $L(x^o,\lambda^o) \geq L(x^o,\lambda)$, for all $\lambda \geq 0$

The *dual function* is

$$h(\lambda) = \min_{x \epsilon X} L(x,\lambda)$$

and the *dual problem* is

$$\max_{\lambda \epsilon D} h(\lambda)$$

where

$$D = \{\lambda \mid \lambda \geq 0, \ \min_{x \epsilon X} L(x,\lambda) \ \text{exists}\}.$$

Theorem 1

The point (x^o,λ^o), for $x^o \epsilon X$, $\lambda^o \geq 0$, is a saddle point for $L(x,\lambda)$ iff:

(a) x^o minimizes $L(x,\lambda^o)$ over X

(b) $g(x^o) \leq 0$

(c) $\lambda^o g(x^o) = 0$

[Note: $\lambda^o \geq 0$ associated with condition (a) are called *generalized Lagrange multipliers* (GLM)].

proof: (=>)

The first inequality (i) above is equivalent to (a). The second (ii) states that

$$f(x^o) + \lambda^o g(x^o) \geq f(x^o) + \lambda g(x^o)$$

$$\Rightarrow (\lambda - \lambda^o) g(x^o) \leq 0 \Rightarrow g(x^o) \leq 0, \text{ for all } \lambda \geq 0$$

Now, $\lambda_i = 0$, $i = 1,\ldots,m$, $\Rightarrow \lambda^o g(x^o) \geq 0$

But since $\lambda \geq 0$ and $g(x^o) \leq 0 \Rightarrow \lambda^o g(x^o) \leq 0$

then $\lambda^o g(x^o) = 0.$

(<=)

$$L(x^o,\lambda^o) = f(x^o) + \lambda^o g(x^o) = f(x^o)$$

$$L(x^o,\lambda) = f(x^o) + \lambda g(x^o)$$

But since $\lambda g(x^o) \leq 0$, $L(x^o,\lambda) \leq L(x^o,\lambda^o)$

Condition (a) is equivalent to Inequality (ii) ||

Theorem 2

If (x^o,λ^o) is a saddle-point for $L(x,\lambda)$, then x^o solves the primal problem.

proof:

$$f(x^o) + \lambda^o g(x^o) \leq f(x) + \lambda^o g(x), \quad \text{for all } x \in X$$

But $\lambda^o g(x^o) = 0 \Rightarrow f(x^o) \leq f(x) + \lambda^o g(x)$

$\Rightarrow f(x^o) \leq f(x).$ ||

Questions that immediately arise include:

1. Do such GLM vectors $\lambda^o \in D$ exist such that the original primal problem is solved?
2. If they do exist, how can they be determined?

Theorem 3

If the primal problem satisfies:

(i) X convex,

(ii) f and g_i convex, $i = 1,\ldots,m$,

(iii) there exists $\bar{x} \in X$ s.t. $g(\bar{x}) < 0$,

then x^o is a solution to the primal problem <=> there exists $\lambda^o \in D$ such that (x^o,λ^o) is a saddle-point for $L(x,\lambda)$.

proof: (see Karlin [19a] or Lasdon [21]). ||

Therefore, for convex programming problems with certain constraint qualification, the existence of a saddle-point is guaranteed. Such assurance is not automatic for more general nonconvex problems.

Theorem 4

$$h(\lambda) \leq f(x), \text{ for all } x \in X, \text{ for all } \lambda \in D$$

proof:

$$h(\lambda) = \min_{x \in X} (f(x) + \lambda g(x))$$
$$\leq f(x) + \lambda g(x)$$

But, for all $x \in X$ and for all $\lambda \in D$, $\lambda g(x) \leq 0$. Therefore,

$$h(\lambda) \leq f(x), \text{ for all } x \in X, \text{ for all } \lambda \in D. \quad ||$$

A *duality gap* exists if $h(\lambda) < f(x)$, for all $x \in X$, for all $\lambda \in D$. In this case, there exists no optimal GLM vector $\lambda^o \geq 0$ such that the primal problem can be solved.

Theorem 5 (Karlin [19a])

$$\min_{x \in X} \max_{\lambda \in D} L(x,\lambda) = \max_{\lambda \in D} \min_{x \in X} L(x,\lambda)$$

if and only if there exists a saddle-point.

proof: (see Karlin [19a] or Lasdon [21]). ||

Theorem 6

The set D is convex, and $h(\lambda)$ is concave over D.

proof: Let $\lambda_1, \lambda_2 \in D$. For $\alpha \in [0,1]$,

$$\lambda = \alpha\lambda_1 + (1 - \alpha)\lambda_2 \geq 0$$

$$h(\alpha\lambda_1 + (1 - \alpha)\lambda_2) = \min_{x \in X} L(x, (\alpha\lambda_1 + (1 - \alpha)\lambda_2))$$

$$L(x, (\alpha\lambda_1 + (1 - \alpha)\lambda_2) = f(x) + (\alpha\lambda_1 + (1 - \alpha)\lambda_2)g(x)$$

$$= \alpha f(x) + \alpha\lambda_1 g(x)$$

$$+ (1 - \alpha)\, f(x) + (1 - \alpha)\, \lambda_2 g(x)$$

Therefore,

$$h(\alpha\lambda_1 + (1 - \alpha)\lambda_2) = \min_{x \in X} [\alpha L(x,\lambda_1) + (1 - \alpha)\, L(x,\lambda_2)]$$

$$\geq \alpha \min_{x \in X} L(x,\lambda_1) + (1 - \alpha) \min_{x \in X} L(x,\lambda_2)$$

$$= \alpha h(\lambda_1) + (1 - \alpha)\, h(\lambda_2) > -\infty. \quad ||$$

Having established the concavity of the dual function, even though f and g may be nonconvex, it is clear that any solution to the dual problem must be a global answer.

Theorem 4 states that this answer will be at least a lower bound for $f(x)$. Unless strict concavity can be established, there may not be a unique λ^o associated with the global solution. In addition, it is important that $h(\lambda)$ be differentiable over the entire set D if gradient-type methods are to be used for converging to the solution of the dual problem. The concept of *g-uniqueness* is used to establish the differentiability of $h(\lambda)$.

Definition: Let $D \subset E^m$ and $h: D \to \bar{E}^1$ be concave over the convex set D, where $\bar{E}^1 = E^1 \cup \{+\infty\} \cup \{-\infty\}$. A vector $\bar{c} \in E^m$ is called a *subgradient* of $h(\cdot)$ at $\bar{\lambda} \in D$ if $h(\lambda) \leq h(\bar{\lambda}) + \bar{c}(\lambda - \bar{\lambda})$, for all $\lambda \in D$. The set of subgradients of $h(\cdot)$ at $\bar{\lambda}$ is represented by $\partial h(\bar{\lambda})$. [Note: the inequality is reversed if h is convex].

Certain properties of $h(\cdot)$ over D can be listed which follow directly from its concavity (see Rockafeller [25a]):

(i) $h(\cdot)$ is continuous on the interior of D.

(ii) $h(\lambda) = +\infty$ for some $\lambda \in \text{int}(D) \Rightarrow h(\lambda) = +\infty$ for all $\lambda \in \text{int}(D)$.

(iii) $h(\cdot)$ is differentiable at $\bar{\lambda} \in D \Rightarrow h(\cdot)$ has a unique subgradient at $\bar{\lambda}$.

Definition: A nonempty set $A \subset E^n$ is called *g-unique* if a map $g(\cdot)$ is constant over A.

Let

$$X(\lambda) = \{x \in X \mid x \text{ minimizes } L(x,\lambda)\}, \text{ for all } \lambda \in D$$

Theorem 7

For any $\bar{\lambda} \in D$ and $\bar{x} \in X(\bar{\lambda})$, $g(\bar{x})$ is a subgradient of $h(\cdot)$ at $\bar{\lambda}$.

proof:

Since $\bar{x} \in X(\bar{\lambda})$,

$$f(\bar{x}) + \bar{\lambda} g(\bar{x}) = h(\bar{\lambda}), \text{ and}$$

$$f(\bar{x}) + \lambda g(\bar{x}) \geq \min_{x \in X(\lambda)} [f(x) + \lambda g(x)] = h(\lambda)$$

(for all $\lambda \in D$)

Subtracting the first line from the second gives

$$\lambda g(\bar{x}) - \bar{\lambda} g(\bar{x}) \geq h(\lambda) - h(\bar{\lambda})$$

or

$$h(\lambda) \leq h(\bar{\lambda}) + g(\bar{x})(\lambda - \bar{\lambda})$$

Therefore

$$g(\bar{x}) \in \partial h(\bar{\lambda}). \qquad ||$$

Corollary 8

If $X(\lambda)$ is g-unique for all $\lambda \in \text{int}(D)$, then $h(\cdot)$ has a unique subgradient, and is therefore differentiable at all points in $\text{int}(D)$.

proof: Follows immediately from concavity of $h(\cdot)$ and Theorem 7. ||

Suppose

(i) X is a closed and bounded subset of E^n.

(ii) $f(x)$ and $g_i(x)$ $(i = 1,\ldots,m)$ are continuous on X

Then $D = (E^m)^+$ since $\min_{x\varepsilon X} L(x,\lambda)$ is guaranteed to exist. (See Luenberger [21a], p. 128).

Theorem 9

If (i) and (ii) above hold, then $X(\lambda)$ is g-unique for all $\lambda\varepsilon D$ iff $h(\cdot)$ is differentiable over D.

proof: (see Lasdon [21]) ||

Notice that a special case of g-uniqueness occurs when there is a unique solution to $\min_{x\varepsilon X} L(x,\bar{\lambda})$ for some $\bar{\lambda}\varepsilon D$, or $X(\bar{\lambda})$ contains only one vector $x(\bar{\lambda})$. In general, then, when g-uniqueness holds at some $\bar{\lambda}\varepsilon D$,

$$\frac{\partial h(\bar{\lambda})}{\partial \lambda_i} = g_i(x(\bar{\lambda})),\ i = 1,\ldots,m$$

Theorem 10

Assume (i) and (ii) above hold, let λ^o solve the dual problem, and assume that h is differentiable at λ^o. Then any element $x^o\varepsilon X(\lambda^o)$ solves the primal problem.

proof:

Since $D = (E^m)^+$ and h is differentiable at λ^o, then the following conditions hold:

$$\text{(a)}\quad \lambda_i^o > 0 \Rightarrow \frac{\partial h(\lambda^o)}{\partial \lambda_i} = g_i(x^o) = 0$$

$$\text{(b)}\quad \lambda_i^o = 0 \Rightarrow \frac{\partial h(\lambda^o)}{\partial \lambda_i} = g_i(x^o) \leq 0$$

Therefore, all the conditions associated with Theorem 2 hold. ||