1972

The computation of optimal growth in economic models

Elmo A. Keller Jr.
Iowa State University

72-19,988

KELLER, Jr., Elmo A., 1934-
THE COMPUTATION OF OPTIMAL GROWTH IN ECONOMIC MODELS.

Iowa State University, Ph.D., 1972
Economics, theory

University Microfilms, A XEROX Company, Ann Arbor, Michigan

The computation of optimal growth in economic models

by

Elmo A. Keller Jr.

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Economics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1972

TABLE OF CONTENTS

		Page
I.	INTRODUCTION TO THE NUMERICAL SOLUTION OF ECONOMIC GROWTH MODELS	1
II.	NUMERICAL PROCEDURES FOR SOLVING OPTIMAL CONTROL PROBLEMS	30
	A. Introduction	30
	B. Gradient Method in Function Space	36
	C. Conjugate Gradient Procedure	43
	D. Davidon Method	54
	E. Other Aspects of Computing	64
	1. Discrete control problems	64
	2. Discrete control growth model with inequality constraints	68
III.	ECONOMIC APPLICATIONS AND NUMERICAL SOLUTIONS	76
	A. Introduction to the Control Problem Computation	76
	B. One-Sector Neoclassical Growth and Optimal Growth Models	79
	C. Two-Sector Optimal Growth Models	116
IV.	GENERALIZATION OF THE COMPUTING FRAMEWROK	143
	A. Generalization of Two-Sector Growth Models	143
	B. Simulation of Optimal Trajectories	156
V.	SUMMARY AND FURTHER RESEARCH	171
VI.	BIBLIOGRAPHY	177
VII.	ACKNOWLEDGMENTS	184
VIII.	APPENDIX A	185
IX.	APPENDIX B	198

I. INTRODUCTION TO THE NUMERICAL SOLUTION OF ECONOMIC GROWTH MODELS

During the past three decades much interest has been directed towards problems of decision making in physical, economic or organizational systems. This interest has been motivated primarily by the important economic benefits which result from correct decisions concerning the allocation and distribution of costly, limited resources. Also it has been inspired by the repeated demonstration that such models can be realistically formulated and mathematically analyzed to obtain good decisions. A third reason for this trend is the arrival of high-speed digital computers which play such an important role in the development of large systems and the coupling of previously separate systems, thereby resulting in decision and control problems of increased complexity. The computer has rendered certain techniques obsolete while making other previously impractical methods feasible and efficient.

Let us examine what is meant by the concept of "best" or "optimal" decision. An approach one may use is where a single, real valued functional summarizing the performance or value of a decision, is isolated and optimized (either maximized or minimized depending on the model), by proper selection among available alternatives. The resulting optimal vector is taken to be the solution to the decision problem.

A transformation from a vector space X into the space of real or complex scalars is said to be a functional on X. Real valued functionals are of direct interest to optimization theory since optimization consists of selecting a vector from a given space to minimize or maximize a given functional.

To facilitate communication in formulating the problem we can classify models into four mutually nonexclusive classes.

(1) Deterministic Model--neither the exogenous variables (determined outside the system), nor the endogenous variables (determined within the model), nor the parameters of the model are allowed to be random variables.

(2) Stochastic Models--at least one of the operating characteristics of the model is a probability density function.

(3) Static Models--neither the variables of the model nor the parameters take time into account.

(4) Dynamic Models--deal with time varying interaction of variables and/or the parameters of the model.

The equations describing the decision model may be deterministic or stochastic, and may be complicated from a mathematical point of view. However, the performance index has a simple underlying structure.

$$\underset{\{x\}}{\text{optimize}} \quad V = f(x,y,a) \quad ,$$

$$x \in X,\ y \in Y \text{ and } a \in A \quad .$$

$$f:\ X \times Y \times A \to R \quad .$$

x = Vector that can be controlled and affects V.

y = Vector that cannot be controlled and affects V.

a = Vector of parameters that affects V.

f = Real valued functional.

The problem then is to find values of the controlled variables which optimize the performance index subject to the restrictions given.

A solution may be obtained either by (1) mathematical analysis, (2) numerical approximation, or (3) conducting experiments on the model (simulation).

This approach of formulating decision problems has the virtue of simplicity and precision but it also has the limitation due to the necessity of selecting a single objective by which to measure results.

Let us now focus on the specific problem of interest, that is intertemporal optimization. Here we have the general problem of choosing functions from function spaces that will optimize a given functional and also satisfy differential or difference equations, initial and/or boundary conditions and possibly other constraints.

Much of the classical theory of dynamic or intertemporal optimization was motivated by problems in physics and in the calculus of variations (47). Associated with these results are mathematicians, Gauss, Euler, Lagrange and others. Much of the early work was in obtaining necessary conditions for the solution of the problem. This approach of the classical calculus of variation was to transform the given optimization problem into another problem, namely, the Euler-Lagrange equation. The function that optimizes the functional also satisfies the Euler-Lagrange equation. However, in most cases the Euler-Lagrange equation turned out to be a nonlinear two point boundary value problem. For a large system this by itself is a highly formidable mathematical problem.

During recent developments of optimization in decision problems, the classical methods have been re-examined, extended sometimes rediscovered and applied to problems having quite different origin than those responsible for earlier development. Some illustrations of these applications would be in optimal economic growth models. For example, in any economic unit choices must be made between provisions for the present (consumption) and provisions for the future (capital accumulation). While more consumption is preferable to less at any moment in time, more consumption means less capital accumulation. The smaller the capital accumulation, the smaller the future output of the economic unit and therefore the smaller future potential consumption. Thus a choice must be made between alternative consumption policies. At one extreme is the policy of consuming as much today even though the potential for future consumption is jeopardized. At the other extreme is the policy of consuming only a subsistence today so as to increase capital and the potential for future consumption. The choices made over time between consumption and capital accumulation generate a set of time paths for consumption, capital, and output for the economic unit. Many growth paths are possible and to choose one of them, one must select an appropriate index of performance for the unit in question. Once this judgment has been made, one faces the problem of choosing an optimal feasible growth path, that is the problem of optimal economic growth. This problem can be considered as a problem of intertemporal optimization.

The solution to the problem is not simple and perhaps cannot be attained even if one defines and finds it. Yet it does seem helpful to have a clear picture of the optimal time path as a guide to the directions in which policies may be modified. Given the technical possibilities, the planners can by varying the time path of investments, vary the time path of consumption per capita.

As Arrow and Kurz (3) and Uzawa (72) point out in an economy that is not centrally planned, the problem of optimal economic growth is that of choosing appropriate mixtures of existing policy controls, such as monetary and fiscal policy, to attain the desired objective. More will be said about this idea in a later section

If the planners have a quantitative and unambiguous set of valuations of the time path of consumption, then by comparing the integral of valuations for any situation, they can obtain a measure of which is better. Hence the problem is to formulate and solve an optimal control problem

The problem can be formulated by taking the state of the system by some state vector x(t) and taking the evolution of the economic system with respect to time by the dynamic equations,

$$\dot{x} = f(x,u,t) \quad , \quad x(t_o) = x_o \quad , \quad h(x(t_f)) = 0 \quad .$$

$$u \in V \quad , \quad \dot{x} = \frac{dx}{dt} \quad . \qquad (1.1)$$

$$x \in R^n \quad , \quad u \in R^r \quad , \quad f : R^{n+r+1} \to R^n \qquad (1.2)$$

where u(t) is a vector of controls or instruments and V a set of admissible controls. Each state vector is assumed to be a continuous function of time, so the trajectory

$$\{x(t)\} = \{x(t) \in R^n \mid t_o \leq t \leq t_f\}$$

is a continuous vector function of time. At any time t in the relevant interval, the choices to be made are characterized by r real numbers $u_1(t)$, $u_2(t)$, ..., $u_r(t)$ called control variables and summarized by the control vector. Each control variable is required to be a piecewise continuous function of time so the control vector

$$\{u(t)\} = \{u(t) \varepsilon R^r \mid t_o \leq t \leq t_f\}$$

is a piecewise continuous vector valued function. The control variables may be chosen subject to certain constraints on their possible values, summarized by the restriction that the control vector at all times in the relevant interval must belong to a given nonempty subset of R^r: $u(t) \varepsilon V$, $t_o \leq t \leq t_f$.

One can take the preferences of those in the system to form the integral performance functional,

$$E = \int_{t_o}^{t_f} L(x,u,t)dt + \emptyset(x(t_f)) \quad . \tag{1.3}$$

Where L is a utility function and Ø is a terminal "bequest function" and $[t_o,t_f]$ represents the planning horizon. Given this general structure (in either discrete or continuous form) one readily observes that the selection of an appropriate economic policy to maximize E is precisely a problem in optimal control.

An example would be where a centrally organized decision making body select patterns of production and investment which would generate a set of time paths of sectoral growth to optimize an index of welfare for the complete economic unit. This type of problem determines the optimal allocation of investment between sectors at any point in time and the optimal time path of consumption.

One can further classify growth models into two principal classes.

1. Consistency models--models by which one attempts to choose a pattern of resource allocation among various sectors of the economy which is consistent with a given set of "targets" given for the end of the planning horizon.

2. Optimal models--models designed to find the best by means of optimizing a utility functional of different time paths of resource allocation over the planning horizon.

In the optimal control models, one has the following elements: (a) performance functional, (b) a dynamic model containing some variables appearing in the criterion functional and (c) a subset of the model variables which can be controlled.

In solving control problems of optimal economic growth one has the goal of finding a decision rule for determining the present control decision subject to certain constraints that will either minimize the deviation from some ideal behavior or that will maximize the functional consisting of a utility function of certain system variables. The performance index is important because it, to a large measure, determines the nature of the resulting optimal control vector. It is highly desirable that the index of performance originate not from a mathematical, but from an applicational point of view. However in certain cases this choice involves compromises between a meaningful evaluation of the system and a tractable mathematical problem.

Difficulties in choosing or constructing an aggregate utility function are recognized. These same problems also apply in obtaining a performance functional. Let us assume that we can construct a collection of such utility functions each possessing various properties and proceed from that point to determine sensitivity measures of how the control policy changes with respect to changes in structural features of the model.

The nature and difficulty of the mentioned control problems vary considerably, depending upon the kinds of information, and structure available in the following interrelated categories:

1. Performance index, initial state, desired final state of the model and the planning time horizon.

2. Characteristics and structural features of the dynamical system equations of the model.

3. Characteristics of the allowable control and state

vectors and the nature of the constraints on them.

4. Permissible interaction between controls and the system equations and the solutions.

In the past decade the theory of optimal control has obtained theoretical tools such as Pontryagin's Maximum Principle (58) (called minimum principle by many American authors; 47, 12) dynamic programming (5) and such numerical techniques as will be discussed later to approximate the solution to optimal control problems, given the necessary information for the model.

The question of how the optimal state and control trajectories change with respect to changes in certain features of the problem when one or more parts of key structural information takes on various values is studied later.

The analysis is in the form of numerical experimentation dealing with nonlinear models under various economic hypotheses about the models. The solutions obtained will be numerical trajectories computed using recently developed numerical algorithms to solve optimal control problems.

Frank Ramsey (60) considered a neoclassical model of production where the optimal trajectory of capital accumulation maximizes the integral over time of the utilities of per capita consumption.

Extensions or elaborations of neoclassical optimal growth models have been compiled by Karl Shell (67) and an extensive bibliography given by Dobell (21), Burmeister and Dobell (13). Treatment is also given by Arrow and Kurz (3) and Uzawa (72). Computation of the optimal paths was not the objective of Shell and associates, Arrow and Kurz or Uzawa. Rather they were concerned with the qualitative analysis of the solution by using Pontryagin's Maximum Principle (58). They analyzed conditions for the existence of optimal control time paths and the asymptotic properties of such paths. They do not specify any computing sequence or procedures to numerically solve the optimal growth models considered, but rather analyze the steady state solutions and their economic meanings.

These theoretical models were designed to analyze the characteristics of any economy in asymptotic optimal growth. One of my objectives is to develop and solve numerically certain finite horizon optimizing growth models which allow the specification of production and welfare relationships in a nonlinear form and thereby analyze some properties of such models.

Deterministic optimal growth models may be divided into two groups: aggregative and disaggregative. The aggregative models are generally based on the assumption of a single sector. The disaggregative models seek to specify the relative rates of growth for several interdependent sectors of an economy. An intermediate case between the two types of growth models is provided by intersectoral models, where the different sectors are completely independent or substantially so.

There exists a wide variety of economic growth models having different degrees of disaggregation, different levels of dynamic relationships and possibly different policy implications. The choice between alternative models presents a difficult task, especially if one's interest is in applying some of the current growth theory to planning and development. Some critical areas include the various linkages which exist between an aggregate model and its disaggregated version and also the implication of certain types of balanced growth which may have an oscillatory tendency.

In Chapter 16 of his book, Morishima (52) treats a model and conditions required for the simultaneous optimization of capital accumulation and population growth. He emphasizes among other things the potential danger of cyclical oscillation in per capita consumption and output in turnpike models of long-run economic growth.

Operational planning models based on specific policy formulations relating to economic growth and stabilization for development planning have been considered by various authors. One may mention the planning and programming models for countries such as the Netherlands (14, 15), Norway (7), at different levels of formulation and actual application and other models formalized by Chenery and Bruno (17), Lange (42), Mahalanobis (48), Klein (40, 41) and others.

Generally development planning models are concerned with economic growth within a medium or long range time horizon. Policy models are usually formulated for short term stabilization purposes, often within a growth framework.

Planning methods specify the role of planning in achieving economic objectives. Planning models without any control are generally either a linear or nonlinear programming problem within an input-output framework. A number of things are required to be fixed in the model and hence their effects on the optimal solution cannot be determined. Some of these include the time horizon and terminal constraints on various state variables. There is no possibility of updating and modifying the solution and no test of sensitivity in a complete sense.

Planning with control as in a control problem using penalty functions to handle terminal constraints allows one to adjust the time horizons, the terminal constraints and to determine sensitivity measures on these parts of the model. One can solve such control problems numerically that may be analytically intractable in terms of elementary functions.

One can mention some reasons for giving thought to dynamic economic models in a frame of reference of optimal control theory. In a centrally-planned economy the planners have a direct influence on the time paths and character of economic growth and may wish to have the benefit of economic analysis in wielding that influence. Also one needs a reference to which other possible time paths may be compared. It may be useful to have as that bench mark an optimal path with respect to certain indices of performance. In addition, one notes that in the individual enterprise economies, the main determinant of savings and hence investment are the desires of business firms to control their survival and growth by internal financing (accumulation of capital), and the concern of individuals with their support in old age and with the economic opportunities of their children. Even in these economies, governments have a considerable influence on savings and on other aspects of economic growth. Thus the same consideration as first mentioned applies there also. Distribution problems between individuals living at the same time are ignored.

Many feasible time paths are possible in a growth problem. One way to compare such feasible paths is to construct appropriate measures of performance and frame the problem in a control problem format. Also comparisons can be made in terms of the squared deviation from a given desired trajectory subject to the constraints of the dynamics of the model. Computation of optimal control also allows one to examine possible feedback relations where the control vector is a function of the state vector over a given horizon.

A second reason is the success of control theory in fields other than economics (4, 12). In the last few years many new algorithms have been developed for the computation of control problems (44, 70, 1, 53). Insight gained using these algorithms on small economic models may give insight to the analysis of larger more complex models and to the feasibility of using such algorithms on large planning models. A comparison of the optimal time paths to the trajectories generated using feedback and simulation is also needed. Computation of small economic models serves to reveal sensitive parameters in the model.

Consider the economic meaning of the Lagrange multipliers in the general classical programming problem.

$$\text{maximize } F(x) \quad , \quad \text{subject to } g(x) = b \ , \tag{1.4}$$

where

$$L(x,\lambda) = F(x) + \lambda'[b-g(x)] \ . \tag{1.5}$$

$x \in R^n \ , \quad g \in R^m$

$n > m$

$F : R^n \to R$

$g : R^n \to R^m$

$b \in R^m \ .$

The Lagrange multipliers at the optimal solution measure the sensitivity of the optimal values of the objective function $F^* = F(x^*)$ to variations in the constraint constants b.

$$\lambda_i^* = \frac{\partial F^*}{\partial b_i} \qquad i=1,2,\ldots,m \quad . \qquad (1.6)$$

For example, if any Lagrange multiplier were equal to zero at the optimal solution, then small changes in the corresponding constraint constant would not affect the optimal values of the objective function. For problems of economic allocation in which the objective function has the dimension of value, and the constraints specify a certain value for a given quantity, the Lagrange multiplier measures the sensitivity of a value to changes in a quantity and hence a price, often called a shadow price.

Given the nonlinear programming problem,

$$\max F(x) \text{ subject to } g(x) \leq b \quad , \quad x \geq 0 \quad . \qquad (1.7)$$

The Lagrange multipliers can be interpreted as in the the classical programming problem,

$$\lambda_i^* = \frac{\partial F^*}{\partial b_i} \qquad i=1,\ldots,m. \qquad (1.8)$$

To the extent that objective functional has the dimension of an economic value and the state variable has the dimension of an economic quantity, then the adjoint variable in a control problem has the dimension of a price, a shadow price. This interpretation of the adjoint variable is the dynamic analogue to the interpretation of the Lagrange multiplier for static optimization problems (36).

The implication of a time path of shadow prices in the control problem and the indirect control through price guidance is a topic of interest. That the market place solves the economic problem of equating supply and demand by successive approximation using feedback to the equilibrating price or prices is a familiar concept. In a single market, each approximation results in naming a price and calculating the difference between supply and demand at that price. The next approximation involves adjusting the previous trial price in a manner governed by the difference, with the idea of causing the difference to vanish.

Lange (43) points out an important limitation of the market is that it treats the accounting problems only in static terms. It does not provide a sufficient foundation for the solution of growth and development problems. In particular it does not provide an adequate basis for long-term planning. For planning economic development, long-term investments have to be taken out of the market mechanism and based on judgment of developmental economic policy. This is because present prices reflect present data, where as investment changes data by creating new technical conditions for production and frequently also by creating new wants. Investment changes the conditions for supply and demand which determine equilibrium prices.

The theory and practice of time staged mathematical programming makes it possible to introduce economic accounting into this process. After setting up an objective function and certain constraints, future shadow prices can be computed. These shadow prices may serve as an instrument of economic accounting in development plans. Actual market equilibrium prices do not suffice here, knowledge of the computed future shadow prices is needed. Here computation does not replace the market, but possibly it may fulfill a function which the market never was able to perform.

Since welfare economics assures us that under certain assumptions (2) as to the utility function and productive process a competitive equilibrium can be identified with an economic optimum, it appears that the method of successive approximations which solves the problem of market equilibrium is also a computational method for solving the problem of optimal resource allocation. An interesting question to consider then would be, "is the reverse true? Does solving the problem of intertemporal optimal resource allocation, generate the time paths of prices?" Certainly for a large economic system a completely centralized organization would require storage capacity and processing that exceeds anything likely to be available. Thus a reason for the computation of optimal economic growth models would be to analyze and study the above question for small economic systems. If the adjoint variables are prices, how does the market mechanism and a central planning process compare with regard to respective transaction costs and iteration costs, information processing and computation. Is the convergence of the control theory algorithm analogous to the convergence of the market mechanism? Some of these questions may be studied by the computation of small models.

For a growth model to be considered operational it must explain the observed process of growth of an economy by means of a set of quantitative variables so that the empirical realism of the model may be tested. In addition it must contain a set of variables amenable to control by one or a set of policy makers such that the observed process of growth may be influenced by the control variables to converge to an optimal process of growth when the optimality condition is defined in some meaningful economic sense. An economic meaning of the computation of growth models would be to show that the model is operational. This would also give an indication of the feasibility of handling large complex models through the same procedure--that of system analysis applied to planning models in a frame of reference of optimal control.

Numerical analysis considerations of the meaning of the computation would include: 1. determination of a numerical solution which could not have been obtained analytically, 2. comparison of the effectiveness of the control algorithms used, 3. indication of the convergence properties of the control algorithms and an approximate measure of their cost, 4. indication of the sensitivity to errors in the computational process by the algorithms considered, 5. comparison of how the algorithms respond to penalty function formulation to handle terminal constraints on the state variables.

As indicated Shell (67) did not, as mentioned, perform computation on the models that were studied. They were concerned with qualitative analysis only. My objective is to consider two problems that were not there explored.

First I consider the problem of how to actually perform the numerical computation of such models. The difficulties of the procedures involved in this computation, and the feasibility of using control theory algorithms to solve such economic models is treated. Also considered is the treatment of what time paths can be realized with respect to various parameter settings of the model. In addition the effect of using penalty functions to handle terminal constraints is studied. In the computational procedure one can obtain an approximate cost measure on the algorithm in terms of the number of iterations required to converge to a satisfactory solution and the amount of computation per iteration. I utilize the conjugate gradient algorithm and the Davidon algorithm. Penalty functions are used to handle terminal constraints on the state variables. By using different algorithms in the control problem computation, one obtains an indication of which algorithm performs better. Speed of convergence to the optimal solution is generally dependent upon the algorithm used.

When one applies these models, certain empirical problems need to be considered. Among these are some of the following. Policy makers and planners have certain preferences which generate various desired values of the control variables. For example as treated in Ijiri (35), planners may consider planning as the process of decomposing given economic goals into a set of subgoals which are more operational and controllable than the main goals. The planning process then becomes one which is directed towards deriving a set of subgoals that will collectively achieve the given goals. The central problem becomes: How does one measure performance in the subunit or subgoal to determine performance in terms of a given goal?

In traditional theory a planner is supposed to be an optimizer. This should not be completely equated to the idea of "rationality" since requisite degrees of knowledge may be absent, for instance when uncertainty is present. It may then be "rational" to be a "satisficier" (49) and thereby proceed "rationally" towards goals that he sets for himself or others rather than seek an optimum.

Suppose that one has a subunit model within a given control or noncontrol framework. This submodel problem is solved to obtain an optimal control vector $u^*(t)$ for the subunit. One could then use this $u^*(t)$ as a desired time path in the aggregative complete model, where the index of performance is the squared deviation from the subunit optimal path and the dynamics reflect the complete model constraints.

$$\text{minimize } J_1 = \int_o^{t_f} (u - u^*)^2 dt \quad . \tag{1.9}$$

$$\dot{x} = g(x,u,t) \quad . \tag{1.10}$$

$$x(o) = x_o \quad , \quad x(t_f) = x_f \quad . \tag{1.11}$$

This suboptimization procedure allows a compromise result to obtain a solution close to the desired path. If the index of performance is larger then a given tolerance, one may sacrifice some in the subunit and modify the subunit control time path $u^*(t)$ and then repeat the process.

This procedure allows for trade offs between the complete control model and subunit models. Computation of such a problem is undertaken in Chapter 3B and 3C to provide insight into the feasibility of such a decomposition procedure.

Both deterministic and stochastic simulation are also utilized as well as the control problem algorithm approach.

Other ways in which the desired path may be determined include treating the desired path as a constant, say consumption per worker subsistence level, with or without a time trend. Sengupta and Walker (66) used as the known desired path a subsistence level of consumption required for the t th year which is a function of the size of the population.

Desired path

$$C^* = C_B \frac{P(t)}{P(B)} \quad ,$$

where P(t) is the population as a function of time, B is a given base year and C_B is the level of consumption for the base year B.

If the objective of the study is a stability analysis and one wishes to have the output of the economic unit maintained close to a desired trajectory, then the desired path may be a predetermined constant level of GNP, again possibly with a time trend. Vanden Bogaard and Theil (73) used the desired output in stability studies, $Y^* = (C_B)(1.5)(1+\alpha)^t$ where α is the presumed net birth rate.

Another possibility for the desired time path would be obtained by not considering the objective functional and eliminating the excess degrees of freedom in the system dynamics by assuming that the control variables are functions of the state variables. The differential system may then be solved and the control variables can be computed and used as desired values for an optimal model. One could begin by assuming simple feedback relations to obtain the desired paths.

Also implicit in applying economic models is the problem of the sensitivity of the model parameters, that is, how the optimal time paths change with respect to changes in the model parameters. Many times the parameters are statistical estimates. Sen (63) and Chakravarty (16) have studied this problem for special one sector models with linear dynamics which admitted analytical solutions. Their investigation was on three main political elements in the formulation of the economic model in terms of maximizing the sum of utilities within a finite horizon; (1) the choice of a utility function, (2) the choice of a time horizon, and (3) the choice of the terminal stock of capital. The first of these is a part of any optimizing program. The latter two result from restricting the period to a finite time horizon. Finite time horizons fit easily into the convenience of planning and the question is not one of a complete break with the future, since the terminal stock of capital provides an adjustable link between the period within the horizon and the period beyond. One could argue that if problem (3) is well solved, the arbitrariness of (2) could be eliminated. The previously mentioned authors studied this problem;

$$\text{maximize } F = \int_{o}^{t_f} e^{-pt} U(C(t))dt \quad . \tag{1.12}$$

$$\text{subject to } \dot{K} = bK(t) - C(t) \quad , \tag{1.13}$$

$$K(o) = K_o \quad \text{and} \quad K(t_f) = K_f \quad . \tag{1.14}$$

K(t) = aggregate stock of capital,

C(t) = aggregate flow of consumption,

b = output-capital ratio

U(C(t)) = utility function = $[C(t) - C^*(t)]^{\alpha}$,

C* = given subsistence level.

The terminal stock K_f is computed from various growth rates g of capital over the time horizon.

Chakravarty (16) using t_f=20, b=1/3, α=.4 concluded by analyzing the various time path data that the best consumption profiles are insensitive to changes in g within the range [.05, .15]. He simply compared the numerical values of the trajectories for different growth rates.

Sen (63) found that for g ε[.15, .325] the consumption profiles are highly sensitive to the growth rate of the capital stock. He defined an over-subsistence consumption function x(t) = C(t) - C*(t) and a "sensitivity indicator", n, as follows:

$$n = \left| \frac{dx(0)}{dg} \Big/ \frac{x(0)}{g} \right| .$$

This indicator was used to analyze the sensitivity of the trajectories.

Another reason for the consideration of the computational aspect of the optimal growth economic model would be so that one could analyze models which have nonlinearities in the system dynamics and also time varying production functions. Sensitivity studies as mentioned above could then be analyzed on nonlinear, time varying problems.

Problems of control are associated with dynamic systems evolving in time. Control or guidance refers to directed influence on a dynamic system to achieve a desired performance. A small number of interesting, nonlinear dynamic optimization problems can be completely resolved analytically by using techniques of 1. Calculus of Variations (32) or 2. Pontryagin's Maximum Principle (58). However, the great majority of dynamic optimization problems must ultimately be solved by computer methods. The reason for this is not that the necessary conditions for optimality are difficult to derive, but rather that the solution of the resulting nonlinear equations is usually beyond analytic tractability.

There are two basic approaches for resolving complex dynamic optimization problems by numerical techniques:

1. Formulate the necessary conditions describing the optimal solution and solve these equations numerically usually by some iterative scheme.
2. Bypass the formulation of the necessary conditions and implement a direct search for the optimum.

Although the field of optimal control has received much specialized attention in recent years, it cannot be disassociated from the noncontrol branches of optimization such as linear programming, nonlinear programming, and the calculus of variations. These noncontrol branches of optimization theory have contributed greatly to the development of iterative techniques for solving the control problem. The terms direct and indirect are frequently used to classify the many numerical techniques that have been used. Indirect methods are those which attempt to produce the optimal control by satisfying the necessary conditions for optimality obtained from the calculus of variations or from Pontryagin's Maximum Principle. In general, the application of these necessary conditions leads to a two-point boundary value problem. Most indirect methods, as a result, are characterized by an iterative modification of either the boundary conditions or the differential equations.

In contrast, direct methods are those that select successive trial control functions based on information obtained from the value of the functional and its gradient for previous control choices. The methods usually require the choice of an initial control function which is used to determine a direction of search in the space of allowable control functions. The control change is the product of the direction of search vector and a scalar called the search direction parameter or search direction stepsize. From the control function, a new direction of search is determined, and the process is repeated. The various direct methods differ mainly in the means used to determine the successive directions of search and the magnitude of the control stepsize taken in those directions. The conjugate direction methods are direct solution algorithms.

The class of numerical techniques called conjugate direction methods combine the computational simplicity of the gradient techniques with the rapid convergence properties typical of second-order techniques. These methods do not require the computation of second-order partial derivatives in determining the direction of search. The improved direction of search results from the assumption that the objective function can be approximated by a quadratic function in the neighborhood of the current search point. The properties of the quadratic function are used implicitly in the derivation of the methods to produce directions of search that are superior to the negative gradient directions. Two such procedures, the conjugate gradient and the Davidon method, will be discussed in later sections.

In the general case of nonlinear systems with non-quadratic performance criteria, the specification of optimal control requires the solution of 2 n simultaneous first order differential equations for an nth order system with mixed boundary values. It represents a difficult problem in numerical analysis because the coupled equations are usually highly unstable.

In the next chapter some of the basic concepts associated with various numerical control algorithm procedures will be considered. The question of existence and uniqueness of an optimal control in what follows is avoided here, as in most numerical treatments, by assuming that a unique optimal control exists.

II. NUMERICAL PROCEDURES FOR SOLVING OPTIMAL CONTROL PROBLEMS

A. Introduction

The development and use of the numerical methods of mathematical optimization is important to many scientific disciplines. As indicated in Chapter 1, an interest, many times, to the management scientist and economist is to a part of optimization referred to as optimal control. This field has received much interest in recent years, but is integrally associated with other optimization areas such as linear and nonlinear programming and the calculus of variations.

This chapter treats a class of iteration techniques for solving the control problem. First the gradient technique is presented. Steepest descent is perhaps the oldest direct method of minimizing an objective function of several variables. The procedure is based on the principle of choosing a trial solution that lies along the direction of maximum decrease of the objective function from the previous iteration. The question of stepsize in the direction of search is important--very small stepsizes are impractical and inefficient, while large stepsizes lead to convergence problems. Curry (19) suggested that from each point in the search, the negative gradient direction is to be followed by a one dimensional minimization of the objective function to determine the optimal stepsize for the next iteration. With that procedure implemented the gradient method becomes a useful computational method. Bryson and Denham (10, 11) and Kelley (39) extended the use of steepest descent to function spaces. These and other authors have incorporated methods for handling terminal state constraints and certain state space and control variable constraints.

Second order direct methods of solving optimal control problems have been developed by Breakwell, Speyer and Bryson (8) and others. These techniques are extensions of Newton's method for minimizing a function of several variables. A quadratic function of n variables can be minimized in one step if the search direction is taken to be the negative gradient direction premultiplied by the inverse Hessian matrix. If the objective function is globally convex, the inverse Hessian matrix evaluated at the search point gives additional second order information for new search direction that leads to faster convergence than the gradient method. Newton's method gives faster convergence at a cost of the evaluation of the inverse Hessian matrix at each step. In addition, if the Hessian matrix is not positive definite everywhere in the search space, Newton's method may not converge at all. Newton's method extensions to function space will not be treated in this chapter, but they have been used to solve control problems (12, 50).

Second-order methods possess rapid convergence near the minimum, but they require greater computational effort than do first order procedures. In addition for starting values far from the minimum in certain problems, they may not converge at all. Two computational techniques that have the efficienty and computational simplicity of first order methods but exhibit convergence properties approaching those of the second order methods will then be treated in Sections C and D of this chapter. These procedures like the first and second order techniques have their origins in finite dimensional algorithms and are called the conjugate gradient and Davidon method.

Basically, the improved directions of search results from the assumption that the objective function can be approximated by a quadratic function in the neighborhood of the current search point. The properties of the quadratic function are used in the derivation of the methods to produce directions of search superior to the negative gradient directions.

Hestenes and Stiefel (33) published the conjugate gradient method as a technique to solve a system of linear algebraic equations. Fletcher and Reeves (26) used the conjugate gradient procedure to minimize a function of several variables, or equivalently, to solve a set of nonlinear equations. Davidon (20) published another conjugate direction method such that when applied to a quadratic function sequentially constructs a matrix which approaches the inverse Hessian matrix. The directions of search chosen are the negative gradient directions premultiplied by the Davidon weighting matrix. He called the procedure a variable metric method, but now it usually is referred to by his name. Fletcher and Powell (25) improved the original version and published computational results. Many others have written about these algorithms. Beckman (5) for one, presented an explanation of the conjugate gradient method.

As in the case of the gradient method and Newton's method, both the conjugate gradient and Davidon's method have been extended to apply to functionals on a suitable function space. Hayes (31) extended the method in 1954. Mehra and Bryson (51), Lasdon et al. (45), Sinnott and Luenberger (68) have also extended and generalized the conjugate gradient method. Willoughby (71) has published computational results of the conjugate gradient algorithm to certain special problems. Tripathi and Narendra (70), Lasdon (44), Adachi et al. (1) have made extensions of the Davidon algorithm to function spaces. The contributions of many of these authors will be treated in later sections.

After treatment of the conjugate gradient and Davidon algorithm to continuous control problems, some other aspects of computing will be considered, such as discrete versions of the problem and a discrete version of the control problem treated by Bruno (9). His treatment of a linear economic growth model leads to a type of algorithm where one approximates the adjoint variables at the initial time and then improves the estimate by an iteration procedure relating linear programming and the dual linear program.

A serious question that arises in a computer based study is whether to formulate and work with a continuous or a discrete time model. One inevitably has to discretize problems for digital computer solutions. The control problem will be formulated first in continuous time and later in discrete time.

My purpose here is to develop and analyze methods as useful tools for solving the following deterministic continuous optimal control problem.

Consider a dynamical system, described by the system of nonlinear differential equations,

$$\dot{x}(t) = f(x,u,t) \quad , \quad f : R^{n+r+1} \rightarrow R^n \tag{2.1}$$

where $x(t)$ is an $n \times 1$ state vector and $u(t)$ is an $r \times 1$ control vector. A performance index,

$$E = \int_{t_o}^{t_f} L(x,u,t)dt \quad , \tag{2.2}$$

is specified where L is defined as: $L : R^{n+r+1} \rightarrow R$. It is assumed that the time interval $[t_o, t_f]$ is fixed and that the state $x(t_o)$ is specified. In addition the system may have inequality and/or terminal constraints given by:

$$h(x(t_f)) = 0 \qquad g : R^{n+r+1} \rightarrow R^s \tag{2.3}$$

$$g(x,u,t) \geq 0 \quad . \tag{2.4}$$

In certain problems the terminal state constraints will be absent and these will be known as free end problems. One seeks a control $u^*(t)$ such that:

a. $u^*(t)$ and the corresponding trajectory $x^*(t)$ minimize the performance index E, satisfy the differential system (Equation 2.1) and initial condition and,

b. the resulting final state $x^*(t_f)$ satisfies Equation 2.3 (part b may not be present in free end problems),

c. $u^*(t)$ and $x^*(t)$ satisfy Equation 2.4 (part c may not be present in certain problems).

We assume that:

1. $f(x,u,t)$ and $L(x,u,t)$ are continuous functions of their arguments and that the first partial derivatives of f and the first and second partials of L with respect to x and u are continuous and that,

2. a unique solution $u^*(t)$ exists.

In the following I discuss n and r vector functions of time in the Hilbert spaces, $L_2^n[t_o,t_f]$ and $L_2^r[t_o,t_f]$. A Hilbert space is a complete normed linear space equipped with an inner product which induces the norm. The inner product of interest is given by

$$[x(t) \mid y(t)] = \int_{t_o}^{t_f} x^T(t)y(t)dt \tag{2.5}$$

where T indicates the transpose. The notation L_x will denote the row vector of partial derivatives of $L(x,u,t)$. The symbol f_x where f is an n-vector indicates an n x n matrix of partial derivatives as does L_{xx}. The symbol f_u represents an n x r matrix.

B. Gradient Method in Function Space

One of the most reliable methods is to decouple the unstable system, integrate n equations forward in time and n equations backward in time. Then maximize the Hamiltonian function H at each time interval using a gradient of H to improve the estimate of the control vector u(t). This algorithm is good for achieving an approximate solution, but final convergence may be intolerably slow. The notation i = 1(1)n denotes that the index i starts at i = 1 and is incremented by 1 until i = n.

Consider the system of differential equations

$$\dot{x}_i(t) = f_i(x,u,t) \quad , \quad x_i(t_o) = x_{o_i} \quad , \quad i = 1(1)n \tag{2.6}$$

where $f_i : R^{n+r+1} \to R$, where $x(t)$ is the n x 1 state vector and $u(t)$ is the r x 1 control vector and t is the independent variable time. The performance criterion is the integral,

$$E = \int_{t_o}^{t_f} L(x,u,t)dt \quad . \tag{2.7}$$

We now define the Hamiltonian function as follows,

$$H(x,u,p,t) = -L(x,u,t) + \sum_{i=1}^{n} p_i(t)f_i(x,u,t) \quad . \tag{2.8}$$

The adjoint system of equations is specified as,

$$\dot{p}_i(t) = -\frac{\partial H}{\partial x_i}(x,u,p,t) \qquad i = 1(1)n \quad . \tag{2.9}$$

Pontryagin's Maximum Principle (58) provides a necessary condition that a specific control $u^*(t)$ is optimal. It states that a control input $u(t)$ which minimizes the performance criterion E, maximizes the Hamiltonian function H. Rather than providing a direct solution to the optimal control problem, the maximum principle produces the result in terms of the solution of another set of differential equations. By maximizing H a relation between $u(t)$, $p(t)$ and $x(t)$ can be generated. Hence the systems 2.6 and 2.9 can be solved, if the necessary initial condition and boundary condition can be determined. Whether the system

2.9 of differential equations in terms of the auxiliary variables $p = (p_1 \ldots p_n)^T$ can be solved depends upon the existence of initial conditions for the system 2.9. Also the coupling between the state equations and the auxiliary equations affects the ability to solve the differential equation system 2.9. The initial and final conditions are usually known for the state variables, but are often not known for the auxiliary variables. Therefore a two-point boundary value problem may result in solving the system 2.6 and 2.9.

Let us consider first the free end point problem with no inequality constraints for which the boundary values on the adjoint vector p(t) are given as:

$$p_i(t_f) = 0 \quad , \qquad i = 1(1)n \quad . \tag{2.10}$$

The actual algorithm would proceed as follows:

a. Select an initial control time vector as a first estimate of u(t).

b. Numerically integrate the system 2.6 forward from t_o to t_f and store the state vector x(t).

c. Integrate the adjoint system 2.9 in reverse time from t_f to t_o using the boundary condition described by Equation 2.10.

d. At each step of the reverse integration the estimate of u(t) is improved according to

$$u^{(k+1)}(t) = u^{(k)}(t) + \alpha_k \frac{\partial H}{\partial u} (x(t),u^{(k)}(t),p(t),t)$$

$$\alpha_k > 0 \qquad (2.11)$$

in such a manner as to maximize H at all times by a steepest ascent procedure. The constant α_k must be determined by an independent search procedure. E can be calculated for different values of α_k and then a polynomial fit made to determine the value of α which minimizes E to be used in the next iteration.

e. Return to step b and repeat the procedure until a specified convergence criterion on u(t), $\frac{\partial H}{\partial u}$, or E is satisfied.

Several variations of the method can be used. If the problem is not a free end point problem, one can define a penalty function and reformulate the problem such that all of the final state variables are free end problems. The performance index is redefined as,

$$E^* = E + \frac{1}{2} \sum_{i=1}^{n} K_i(x_i(t_f) - \bar{x}_i)^2 \qquad (2.12)$$

where the terminal constraints are

$$x_i(t_f) = \bar{x}_i \qquad i = 1(1)n \quad . \qquad (2.13)$$

A minimum of E* is now determined without requiring the terminal values of the state variables to satisfy constraints 2.13 exactly , but instead to require that a "penalty" be paid for any deviation from the terminal values.

With the mentioned modification, one can then use the previous algorithm with the part c (conditions on $p_i(t)$) replaced by the following conditions:

$$p_i(t_f) = \frac{\partial E^*(t_f)}{\partial x_i} = K_i(x_i(t_f) - \bar{x}_i) \quad . \qquad (2.14)$$

The trajectories x(t) and u(t) which minimize E* are close to the trajectories which minimize E subject to the specified end point conditions.

The principal advantage of the gradient method is that convergence is not contingent upon a good initial estimate of the control trajectory. One is assured that the value of the functional to be minimized is decreased at each succeeding iteration. Some disadvantages are that the convergence, although relatively good in the beginning of the iterative sequence, often deteriorates severely as the optimum trajectory is approached. Also the penalty function method required to solve problems with specified terminal conditions introduces arbitrary constants K_i which are required to be "large" at least for the final iteration. If the constants are chosen too large at any point in the iteration cycle, the new control will tend to improve the specified terminal values without much weight being placed on improving the actual functional to be minimized. If the constants K_i are too small, the terminal conditions will not be satisfied. Thus in practice, the success of the method if there are terminal constraints depends upon judicious choices of the penalty constants K_i. In Section A of Chapter 3 a report is given on my experience with the use of penalty functions for handling terminal state constraints.

At this point I would like to clarify a notational procedure used in certain subsequent sections and by many American authors (47, 12). This involves a slight change in the statement of the necessary conditions for the control problem 2.6-2.7. Multiplying the Hamiltonian Equation 2.8 by (-1), one obtains,

$$-H(x,u,p,t) = L(x,u,t) + \sum_{i=1}^{n} (-p_i(t))f_i(x,u,t) \quad . \tag{2.15}$$

Now by redefining the adjoint variables,

$$\lambda_i(t) = -p_i(t) \qquad i = 1(1)n \quad ,$$

the following relationship is determined,

$$-H(x,u,\lambda,t) = L(x,u,t) + \sum_{i=1}^{n} \lambda_i(t)f_i(x,u,t) \quad . \tag{2.16}$$

Letting $-H(x,u,p,t) = V(x,u,\lambda,t)$, one notes that minimizing V with respect to u is equivalent to maximizing H with respect to u, where H is equal to -V. Also the new differential equations,

$$\dot{\lambda}_i(t) = -\frac{\partial V}{\partial x_i} \qquad i = 1(1)n \quad , \qquad (2.17)$$

are equivalent to the differential Equations 2.9. This is seen as follows:

$$\dot{\lambda}_i(t) = -\frac{\partial L}{\partial x_i} + \sum_{i=1}^{n} (-\lambda_i(t)) \frac{\partial f_i}{\partial x_i} \qquad i = 1(1)n \quad .(2.18)$$

Substituting $\lambda_i(t) = -p_i(t) \qquad i = 1(1)n$

$$-\dot{p}_i(t) = -\frac{\partial L}{\partial x_i} + \sum_{i=1}^{n} p_i(t) \frac{\partial f_i}{\partial x_i} \qquad i = 1(1)n$$

$$\dot{p}_i(t) = \frac{\partial L}{\partial x_i} - \sum_{i=1}^{n} p_i(t) \frac{\partial f_i}{\partial x_i}$$

$$\dot{p}_i(t) = -\frac{\partial H}{\partial x_i} \qquad i = 1(1)n \quad .$$

The necessary conditions for the solution of the control problem become Equations 2.6, 2.17 and minimize $V(x,u,\lambda,t)$ with respect to $u(t)$. This formulation gives rise to the term "minimum principle" rather than "maximum principle". The two are equivalent and in what follows most problems are considered in the format of the minimum principle. The Hamiltonian H of Section A is defined as -H in subsequent sections, the adjoint variables as $\lambda_i(t) = -p_i(t)$. With this note, the notation in following sections should be clear whether it is used in the framework of the minimum principle or the maximum principle.

C. Conjugate Gradient Procedure

This technique is an extension of the Fletcher-Reeves method (26) to control problems. If terminal conditions and inequality constraints are present, the problem must be converted to an unconstrained form, possibly by penalty functions. As in the steepest descent method, the gradient trajectory must be computed and stored. In addition, the conjugate gradient technique requires the computation of the norm of the gradient and the storage of one other trajectory, the actual direction of search. Lasdon _et al._ (45) have shown that the direction of search in the function space generated by the conjugate gradient method are such that the objective functional is decreased at each step.

Like most other iterative methods, this procedure cannot distinguish between local and global minima. In general, the best that can be expected is efficient convergence to the bottom of whatever valley it starts in. The usual procedure for problems with local minima is to rerun the method with different starting values.

We note that the following problem,

$$\text{minimize } J = g(x(t_f)) + \int_{t_o}^{t_f} L(x,u,t)dt \quad , \qquad (2.19)$$

subject to $\dot{x}_i = f_i(x,u,t) \quad x_i(t_o) = x_{o_i} \quad , \quad i = 2(1)n$

(2.20)

can be reformulated as follows. Define a new state variable x_1 such that

$$\dot{x}_1 = L(x,u,t) \quad x_1(t_o) = 0 \quad . \tag{2.21}$$

The index of performance 2.19 can then be rewritten as follows;

$$J = g(x(t_f)) + \int_{t_o}^{t_f} \dot{x}_1 dt = g(x(t_f)) + x_1(t_f) = \emptyset(x(t_f)) \tag{2.22}$$

subject to $\dot{x}_i = f_i(x,u,t) \quad x_i(t_o) = x_{o_i} \quad i = 1(1)n$.

(2.23)

It is assumed that given a control vector u, Equation 2.22 and 2.23 can be solved for a unique state vector $x = x(u)$, and thus $J = J(u)$ is a function of u alone. The index of performance Equation 2.19 or in the alternate form Equation 2.22 may include penalty functions to account for terminal state conditions or other constraints. In what follows let u(t) be a single control function ($r = 1$). The extension to the multicontrol problem is straightforward.

The conjugate gradient algorithm requires the computation of the gradient trajectory. Let H, the Hamiltonian, be defined as:

$$H = \sum_{i=1}^{n} \lambda_i f_i = \lambda_1 L + \sum_{i=2}^{n} f_i \lambda_i \quad , \tag{2.24}$$

and

$$-\dot{\lambda}_i = \sum_{i=1}^{n} \lambda_i \frac{\partial f_i}{\partial x_i} \quad , \tag{2.25}$$

$$\lambda_i(t_f) = \left. \frac{\partial \emptyset}{\partial x_i} \right|_{t=t_f} \quad , \qquad i = 1(1)n \tag{2.26}$$

and the gradient is

$$g(u) = \frac{\partial H}{\partial u} \quad . \tag{2.27}$$

Let $u_i(t)$ be the ith approximation to the optimal control $u^*(t)$. The corresponding gradient $g(u_i)$ is computed by solving the state Equations 2.23 forward with $u = u_i$, solving the adjoint system 2.25 with conditions 2.26 backwards in time and then computing $g(u_i)$ from 2.27.

One then proceeds as follows:

$$u_o = \text{arbitrary} \quad , \tag{2.28}$$

$$g_o = g(u_o) \quad , \tag{2.29}$$

and

$$s_o = -g_o \quad .$$

Choose $\alpha = \alpha_i$ to minimize $J(u_i + \alpha s_i)$ (use an independent search routine to compute α) and then (2.30)

$$u_{i+1} = u_i + \alpha_i s_i \qquad \alpha_i > 0 \tag{2.31}$$

$$g_{i+1} = g(u_{i+1}) \tag{2.32}$$

$$\beta_i = (g_{i+1} \mid g_{i+1}) \Big/ (g_i \mid g_i) \tag{2.33}$$

$$s_{i+1} = -g_{i+1} + \beta_i s_i \tag{2.34}$$

where

$$(g_i \mid g_j) = \int_{t_o}^{t_f} g_i^T(t) g_j(t) dt \quad . \tag{2.35}$$

The new direction of search s_{i+1} is not the negative gradient direction $-g_{i+1}$, but is computed using Equation 2.34. The distance traveled in this direction is determined by the one dimensional search problem of $J(u_i + \alpha s_i)$ in Equation 2.30. One iterates by improving u_i at each step by generating search vector s_i using Equations 2.30 through 2.33 until a convergence criteria is satisfied. Lasdon et al. (45) have shown that if u(t) is an element of a Hilbert space Q and J(u) a Frechet differentiable mapping (47) from Q to the real numbers, then the conjugate gradient method when applied to J(u) generates directions s_i which are always directions of descent,

$$\frac{d}{d\alpha} J(u_i + \alpha s_i)\Big|_{\alpha=0} < 0 \quad . \tag{2.36}$$

In the section on the gradient method the topic of penalty functions was introduced. Penalty functions were used to insure that the terminal conditions on the state vector were satisfied. We now wish to consider the optimal control problem with inequality constraints on the state and/or control vectors. Such problems can be solved numerically by converting them to a sequence of problems without inequalities by means of penalty functions. The type of penalty function most often used takes on small values when the state and control vectors are within the constrained set and increasingly large values when they are outside the set. This approach forces satisfaction of the constraints to a desired tolerance. Such functions have been used by Bryson and Denham (11), McGill (50) and others.

The algorithms treated so far apply to problems in which there are no inequality constraints on the control and/or state variables.

Linear (in both system and index of performance) optimal control problems must have control and/or state constraints to be well posed. For such problems the solution is always on a boundary constraint. For nonlinear problems with state and/or control constraints, part of the solution may be on the constraint boundaries (constrained areas) and part may be inside the constraint boundaries (unconstrained areas).

Integral penalty functions form an alternate approach to treat the above type of problem. Consider the scalar inequality constraint

$$g(x,u,t) \leq 0 \quad \text{for all } t_o \leq t \leq t_f \quad . \tag{2.37}$$

The performance index J, (Equation 2.19) may be augmented

$$J^* = J + \mu[\int_{t_o}^{t_f} [g(x,u,t)]^2 \, H(g)dt] \tag{2.38}$$

where

$$H(g) = 0 \quad \text{if } g \leq 0$$

$$= 1 \quad \text{if } g > 0 \quad . \tag{2.39}$$

By a suitable choice of the constant μ (positive if J is to be minimized and negative if J is to be maximized) the constraint 2.37 can be approximately satisfied. If $|\mu|$ is taken too large the previous iterative algorithm will tend to concentrate more on satisfying the constraint than on maximizing or minimizing the performance index. As a result convergence is slow.

Fiacco and McCormick (22, 23, 24) have extended the penalty function formulation. They have considered penalty functions of the above type and interior penalty functions for nonlinear programming problems. The interior method works from inside the constraint set, with the penalty increasing as the boundary is approached. Hence this method seems to avoid many of the problems associated with the irregularity of the constraint boundary. Lasdon, Waren and Rice (46) have extended the interior penalty function technique to control problems as follows.

Consider the problem formulated in Equations 2.19 through 2.23. Add to that formulation the following two constraints.

$$h(x(t_f)) = 0 \qquad h : R^n \to R^m \qquad m < n \quad . \tag{2.40}$$

$$g(x,u,t) \geq 0 \qquad h : R^{n+r+1} \to R^s \quad . \tag{2.41}$$

Since as mentioned, one assumes that given u(t), Equation 2.20, the differential system and initial condition then yields $x = x(u)$. Hence the constraint Equation 2.41 $g_i(x(u),u,t)$ can be formulated as $g_i(u,t)$ and the objective function in Equation 2.22 $\emptyset(x(t_f))$ as $\emptyset(u)$.

Define the set,

$$G(t) = \{x(t) \mid h(x(t)) = 0\} \quad .$$

Let S denote the set of all controls u which together with their associated state trajectories x satisfy,

(i) $x(t_f) \in G(t_f)$ and $\forall t_o \leq t < t_f \quad x(t) \notin G(t)$,

(ii) $g(x,u,t) \geq 0$ for $t_o \leq t \leq t_f$.

Define S^o as the subset of S for which (ii) above becomes

$g(x,u,t) > 0$ for $t_o \leq t \leq t_f$.

The inequality constrained problem Equations 2.22, 2.23, 2.40 and 2.41 can be converted to a problem without inequality constraints by adding a penalty function to the objective index Equation 2.22. This yields the so-called P-function:

$$P(u,r) = \emptyset(x(t_f)) + r \sum_{i=1}^{s} \int_{t_o}^{t_f} \frac{1}{g_i(x,u,t)} dt \quad , \qquad (2.42)$$

where r is a positive scalar. Choose $r_1 > 0$ and $u_o \in S^o$ and consider the problem of minimizing $P(u,r_1)$ starting from u_o, subject to the differential Equations 2.23 and terminal conditions 2.40. This will be called the P-problem.

If the penalty function term

$$r \sum_{i=1}^{s} \int_{t_o}^{t_f} \frac{1}{g_i} dt \qquad (2.43)$$

approaches infinity as any g_i approaches zero for $t \in [t_o,t_f]$, this then leads one to expect that a relative minimum of $P(u,r_1)$ exists in S^o. Lasdon et al. (46) have shown this to follow since the trajectory of steepest descent of P starting from u_o, a path of which P(u,r) is strictly decreasing, cannot penetrate the boundary of S. The minimizing point depends on the choice of r_1 and is denoted by $u(r_1)$.

Now consider repeating this minimization for a sequence of r values $r_1 > r_2 > r_3 \dots r_k > 0$. Each minimizing point $u(r_k)$ is in S^o. Further, by reducing r, the influence of the penalty function term 2.43 which penalizes closeness to the constraint boundaries, is reduced and in minimizing P more computational effort is concentrated on reducing Ø. Thus the sequence of points $u(r_1)$, $u(r_2)$, ..., $u(r_k)$ can come closer and closer to the boundary of the set S if it is needed and profitable, in terms of reducing Ø. Thus in the limit $r \to 0$ one would expect that the minimizing point u(r) approaches the solution of the inequality constrained problem.

One must restrict u to be in S since P(u,r) may have a minimum exterior to S and only those within S are of interest. In practice one can use minimization techniques which only need account for the terminal constraints 2.40, such as gradient methods.

Lasdon <u>et</u> <u>al</u>. (46) have shown that the sequence of P-minima converges to the minimum objective value of the original problem. If the problem stated by Equation 2.22 is linear in x and u for all t and Equation 2.23 in convex in x and Equation 2.22, $h(x(t_f))$, is linear and each component of Equation 2.23, $g(x,u,t)$, is a concave function of x and u for all t, then the problem is one of minimizing a convex functional over a convex set S in u space. Such a problem has no local minima in S^o distinct from the global minimum. In order to establish the existence of a P-problem minimum in S^o, the following assumptions need to be made:

1. S^o is not empty.
2. $\min_{u \varepsilon S} \emptyset(x(t_f)) = v_o > -\infty$.
3. If there exists $t^* \varepsilon [t_o, t_f]$ such that for some i $g_i(u(t^*), x(t^*), t^*) = 0$, then

   $$\int_{t_o}^{t_f} \frac{1}{g_i}\, dt = \infty \quad .$$

4. The functional $\emptyset$ and all components of the vectors h and g are continuous in u for all $u \varepsilon S$.
5. $\{u \mid (\emptyset < k) \text{ and } u \varepsilon S\}$ is totally bounded for any finite k.

Definition: If u_o is a point in S^o then a local minimum of the function $P(u,r)$ relative to u_o is a point $u(r) \varepsilon S$ with the property that in a small neighborhood of $u(r)$ there is no point in S with a lower value of P, and $P(u(r),r) < P(u_o,r)$. Lasdon et al. (46) have proved the following results about this SUMT application to optimal control problems.

Theorem I.

Any local minimum of $P(u,r_o)$ over the set S relative to $u_o \in S^o$, is finite and at least one such point exists.

Theorem II.

Under the assumptions 1-5

$$\lim_{r_k \to 0} [\min_{u \in S} P(u,r_k)] = v_o \quad .$$

Corollary:

1. $\lim_{r_k \to 0} J(u_k) = v_o$

2. $\lim_{r_k \to 0} r_k \left[\sum_{i=1}^{s} \int_{t_o}^{t_f} \frac{1}{g_i(u_k,x_k,t)} dt = 0 \right] \quad .$

The preceding results do not require any convexity assumptions. It is only necessary that the global minimum of $P(u,r_k)$ for $u \in S^o$ be determined for each r_k.

The solutions to each P-problem (x_k,u_k) satisfies the following conditions:

$$\dot{x} = f(x,u,t) \qquad x(t_o) = x_o \tag{2.44}$$

$$\dot{\lambda} = -H_x + \sum_{i=1}^{s} \frac{r_k}{g_i^2} (g_i)_x \tag{2.45}$$

where $H = \lambda^T f$ and $h(x(t_f)) = 0$

$$\lambda(t_f) = (\emptyset_x + h_x^T v)\Big|_{t=t_f} \tag{2.46}$$

$$H_u - \sum_{i=1}^{s} \frac{r_k}{g_i^2}(g_i)_u = 0 \tag{2.47}$$

where v is a vector of appropriate penalty function constants.

D. Davidon Method

Consider the control problem as formulated in the previous section. The objective function may include penalty functions on the terminal constraints and interior penalty functions if there are inequality constraints. Hence the problem can be framed as an unconstrained control problem and solved by the sequential unconstrained minimization techniques discussed in the previous sections.

Recent results have indicated that the most efficient methods for unconstrained minimization which do not require second derivatives are those which, when applied to a quadratic function, generate conjugate directions (25, 20, 56). Hence a quadratic function of n variables can be minimized in n steps or less. As indicated in the previous section, the conjugate gradient method has been extended to optimal control problems. This section considers a different conjugate direction method, Davidon's algorithm, which appears to be more efficient than the conjugate gradient method (59). Extension of Davidon's method for minimization in n variables to control problems has been made (70, 1, 53). Consider first the Davidon method for minimization of functions. Given a scalar function f of n variables $x_1, x_2, \ldots, x_n$ represented by a vector x, the method can be described as follows.

1. First an arbitrary starting point x^o and a symmetric positive definite matrix H^o (generally the identity matrix) are selected.

2. Knowing x^i, the gradient $g^i = f_x(x^i)$ is computed.

3. $P^o = -H^o g^o$. For the succeeding iterations, the H^i matrix is computed by

$$H^i = H^{i-1} + (\sigma^{i-1} \sigma^{i-1\,T}) / \sigma^{i-1\,T} y^{i-1}$$

$$-(H^{i-1} y^{i-1} y^{i-1\,T} H^{i-1})/(y^{i-1\,T} H^{i-1} y^{i-1})$$

where $\sigma^{i-1} = x^i - x^{i-1}$ and $y^{i-1} = g^i - g^{i-1}$. Then $p^i = -H^i g^i$.

4. The next point x^{i+1} is obtained by a one-dimensional search

$$f(x^{i+1}) = \min_{\alpha} f(x^i + \alpha p^i) \quad .$$

5. Go back to step 2 and repeat until a convergence criterion is satisfied.

The extension to optimal control problem follows (53, 70, 1). The problem considered is that of minimizing a functional

$$J = \int_{t_o}^{t_f} L(x,u,t)dt + \phi(x(t_f)) \qquad (2.48)$$

subject to the state equations,

$$\dot{x} = f(x,u,t) \qquad x \in R^n,\ u \in R^r \qquad r \leq n$$

$$x(t_o) = x_o \quad . \qquad (2.49)$$

If there are inequality or terminal constraints they are handled by a penalty function formulation.

For this problem, for a given u, the gradient of J with respect to u on the constraint surface is given by

$$g = L_u(t) + f_u^T(t)\lambda(t) \quad . \qquad (2.50)$$

The adjoint vector λ satisfies

$$-\dot{\lambda} = f_x^T(t)\lambda(t) + L_x^T \qquad \lambda \in R^n \qquad (2.51)$$

where

$$\lambda(t_f) = \frac{\partial \emptyset}{\partial x}(x(t_f)) \quad .$$

Suppose the terminal state constraints are to be treated via penalty functions so that the augmented performance functional $J = J_1 + \frac{1}{2}\psi^T P\psi$ is to be minimized subject only to the differential constraints. Here, P is a p x p positive definite matrix of penalty constants. Then, if i denotes the iteration number, the algorithm can be stated as follows:

1. For i = 0 choose an initial control vector $u_o(t)$.
2. Integrate the state equations $\dot{x} = f$ from t_o to t_f.
3. Define the Hamiltonian function $H = L + \lambda^T f$ and integrate the adjoint equations

   $$\dot{\lambda} = -\partial H/\partial x, \; \lambda(t_f) = \partial\emptyset/\partial x(t_f) + [(\partial\psi/\partial x(t_f)]^T P\psi \; ,$$

   from t_f to t_o.
4. Compute the gradient vector $g_i = g[u_i(t)] = \partial H/\partial u$.
5. If i > 0 compute the auxiliary functions

   $$y_i(t) = g_i - g_{i-1}$$

   $$z_i(t) = u_i - u_{i-1}$$

   $$a_i(t) = y_i, \; i = 1$$

   $$y_i + \sum_{j=2}^{i}[(b_{j-1}|y_i)b_{j-1} - (c_{j-1}|y_i)c_{j-1}] \; ,$$

   $$i > 1$$

$$b_i(t) = z_i/(z_i|y_i)^{1/2}$$

$$c_i(t) = a_i/(a_i|y_i)^{1/2}$$

where $(v|w)$ denotes the inner product

$$\int_{t_o}^{t_f} V^T w \, dt$$

6. Compute the direction of search

$$p_i(t) = -g_i, \quad i = 0$$

$$-g_i - \sum_{j=1}^{i} [(b_j|g_i)b_j - (c_j|g_i)c_j] \ , \ i > 0 \ .$$

7. Let $u_{i+1}(t) = u_i(t) + \alpha_i p_i(t)$ and determine α_i by performing a one-dimensional minimization of J: $J(u_i + \alpha_i p_i) \leq J(u_i + \gamma p_i)$ for all positive γ.

8. Replace i by i + 1; if i = q, where q is the predetermined restart integer, set i = 0 before returning to step 2.

Observe that step 5 requires that $rN[1 + 2(q - 1)]$ values be stored if a table of N values is used to represent each time function.

In the preceding algorithm, it is noted that as the iterations proceed, the number of vector functions to be stored increases. To remove the difficulty the preceding steps are carried out for only q iterations. Then the procedure is repeated starting with the steepest descent direction, the negative gradient direction, at the (q+1) step.

Pierson and Rajtora (57) have presented additional computational experience with the Tripathi and Narenda version of the Davidon algorithm applied to control problems. They conclude that the algorithm, when applied to nonlinear optimal control problems incorporating penalty functions is at least competitive and probably superior to the conjugate gradient method. My computational experience, which is reported in Chapter 3, totally supports that claim. Also my experience indicates that the restart feature is actually an advantage rather than a practical necessity. In the problems that I considered, q was selected small, say 3, 4 or 5. This makes the storage requirement for the algorithm small and the convergence rate is generally enhanced.

One should note that the search of the Hilbert space of controls for the optimal control is restricted to controls satisfying Equations 2.49 and 2.51. The condition

$$g(u^*(t)) = \frac{\partial H}{\partial u} = 0$$

holds only at the minimum of $J(u(t))$. The expression,

$$g(u(t)) = \frac{\partial H}{\partial u} \quad ,$$

is the gradient to the Hamiltonian and points in the direction of increasing J. This is seen by noting that the first variation in J given from Equation 2.48 is

$$\delta J = \left.\frac{\partial \emptyset}{\partial x}\right|_{t=t_f} \delta x_f + \int_{t_o}^{t_f} \delta L \, dt \quad . \tag{2.52}$$

The notation of δJ represents the first order approximation to $J(\hat{u}(t)) - J(u(t))$ where $\hat{u}$ is a given nominal control. Using the definition of the Hamiltonian $H(x,u,\lambda,t) = L(x,u,t) + \lambda^T f(x,u,t)$ and requiring the satisfaction of the state differential Equation 2.49 results in,

$$\delta L = \delta(H - \lambda^T f) = \delta(H - \lambda^T \dot{x}) \tag{2.53}$$

$$= \frac{\partial H^T}{\partial u} \delta u + \frac{\partial H^T}{\partial x} \delta x - \lambda^T \delta \dot{x} \tag{2.54}$$

or rewriting using Equation 2.52 one obtains,

$$\delta J = \left.\frac{\partial \emptyset}{\partial x}\right|_{t=t_f} \delta x_f + \int_{t_o}^{t_f} [\frac{\partial H^T}{\partial u} \delta u + \frac{\partial H^T}{\partial x} \delta x - \lambda^T \delta \dot{x}] dt \quad . \tag{2.55}$$

Integrating the last term in the integral by parts, where the respective vector components are:

$w_i = \lambda_i$ $\qquad$ $V_i = -x_i$

$$dw_i = \dot{\lambda}_i \qquad\qquad dV_i = -\dot{x}_i dt$$

$$\delta J = \frac{\partial \emptyset}{\partial x}\bigg|_{t_f} \delta x_f - \lambda^T(t)\delta x(t)\bigg|_{t_o}^{t_f} + \int_{t_o}^{t_f} [\frac{\partial H}{\partial u}^T \delta u + \delta^T x(\dot{\lambda} + \frac{\partial H}{\partial x})]dt \ . \tag{2.56}$$

However, $\delta x(t_o) = 0$ because the initial conditions are fixed. Using the optimality conditions 2.51, Equation 2.56 becomes,

$$\delta J = \int_{t_o}^{t_f} (\frac{\partial H}{\partial u})^T \delta u \, dt \quad . \tag{2.57}$$

If the variation of the control u is along a direction of search s then,

$$\delta u = s\delta\alpha \quad , \tag{2.58}$$

where α is the scalar search-parameter. Thus the derivative of J along s is given by the inner product of

$$g = \frac{\partial H}{\partial u}$$

and s,

$$\frac{dJ}{d\alpha} = \int_{t_o}^{t_f} [\frac{\partial H}{\partial u}^T s]dt \quad . \tag{2.59}$$

Therefore $\frac{\partial H}{\partial u} = g(u)$ is analogous to the gradient vector in finite dimensional analysis.

This leads to a discussion of the one dimensional search procedure used to compute the optimal α for each iteration. Utilizing Equation 2.59 to compute $\frac{dJ}{d\alpha}$, the one dimensional minimization procedure is based on using a cubic polynomial fit relating $J(\alpha)$ and α. The functional J is evaluated at least twice and also two values of the derivative $\frac{dJ}{d\alpha}$ using Equation 2.59 are computed, hence a cubic polynomial can be determined. The positive value of α that minimizes the cubic polynomial is then chosen for the stepsize parameter for the next iteration. The procedure is similar to techniques used for finite dimensional problems (26) and will be explained in what follows.

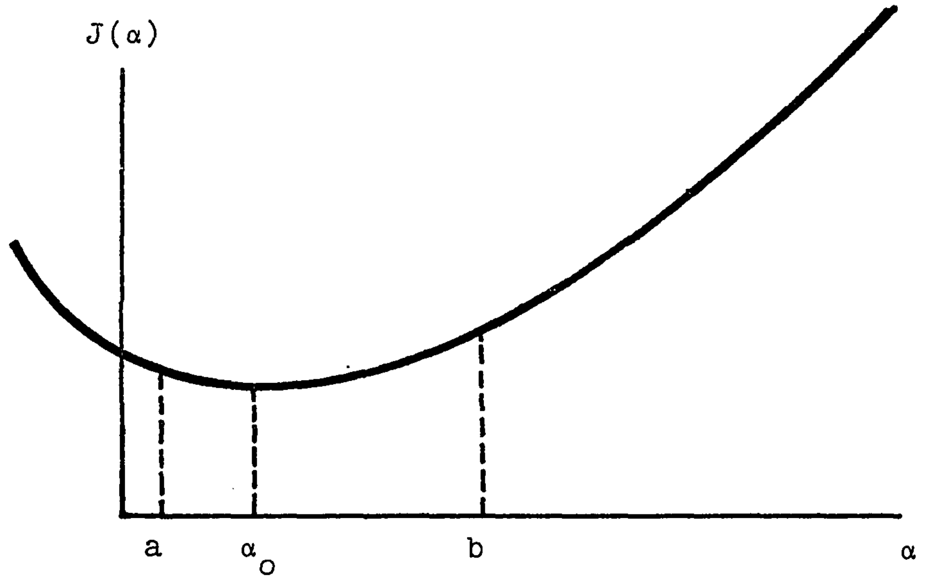

Figure 2.1. One dimensional minimization

The first step estimate for improving the control is given by

$$h = .1/(\int_{t_o}^{t_f} (g^T g)\, dt)^{1/2} \quad .$$

This estimate is used as the initial value to start the procedure at the first iteration. Then $J'(\alpha)$ where the prime indicates derivative with respect to the stepsize parameter α, is examined at the points $\alpha = 0, h, 4h, 16h, \ldots, a, b$. The symbol b represents the first of these values at which $J'(b)$ is nonnegative or $J(b)$ has not decreased. It then follows that α_m is bounded in the interval $a < \alpha_m \leq b$ where α_m is the optimal stepsize parameter to be used in the next iteration.

The next stage uses the cubic interpolation given by Davidon (20) where the positive critical value is computed from the cubic, fitted from the information contained in $J(a)$, $J'(a)$, $J'(b)$, and $J(b)$. One defines

$$Z = \frac{3(J(a) - J(b))}{b - a} + J'(a) + J'(b) \quad , \tag{2.60}$$

$$W = (Z^2 - J'(a)J'(b))^{1/2} \quad , \tag{2.61}$$

then the estimate α_e of α_m is given by,

$$\alpha_e = b - (\frac{J'(b) + W - Z}{J'(b) - J'(a) + 2W})(b - a) \quad .$$

If neither J(a) nor J(b) is less than $J(\alpha_e)$, then α_e is accepted as the estimate of α_m. A check on the value of α_m is the closeness to zero of $J'(\alpha_m)$. If J(a) or J(b) is less than $J(\alpha_e)$, then according as $J'(\alpha_e)$ is positive or negative, the interpolation is repeated over the subinterval (a,α_e) or (α_e,b) respectively. The reinterpolation used here, if the cubic procedure did not work, is a form of linear interpolation on smaller and smaller intervals with an exit after a fixed number of trials.

This technique of choosing the optimal stepsize of the search direction worked well for the applications of both the conjugate gradient and Davidon algorithms.

E. Other Aspects of Computing

1. Discrete control problems

One inevitably must discretize problems for digital computer solution. One can work with a continuous time model and discretize to solve by discrete variable methods or the model can be represented as a discrete multistage system and solved directly.

The mechanics of setting up a discrete time optimal control problem will be described and then solution methods considered. Each of the previously mentioned continuous solution algorithms has a discrete analogue. Also so does the penalty function formulation.

Consider the problem with no inequality constraints of finding the sequence u(0), u(1), ..., u(N-1) and x(1), x(2), ..., x(N) to minimize

$$J = \emptyset(x(N)) + \sum_{i=0}^{N-1} L^i(x(i),u(i)) \quad , \tag{2.62}$$

$$L^i : R^n \times R^r \to R \quad , \qquad \emptyset : R^n \to R$$

subject to the constraints (x is an n vector and u an r vector)

$$x(i+1) = f^i(x(i),u(i)) \qquad i = 0,1,2,\ldots,N-1$$

$$x(0) = x_o \qquad\qquad f^i : R^n \times R^r \to R^n \tag{2.63}$$

$$h(x(N)) = 0 \qquad\qquad h : R^q \to R^n \qquad q \le n \;. \tag{2.64}$$

We can formulate the terminal constraint 2.64 as an exterior quadratic penalty function $K[h(x(N))]^2$ and include it in the $\emptyset(x(N))$ function. Let us adjoin 2.62 with a sequence of multipliers $\lambda(i)$,

$$\bar{J} = \emptyset(x(N)) + \sum_{i=0}^{N-1} \{L^i(x(i),u(i)) + \lambda^T(i+1)$$

$$[f^i(x(i),u(i)) - x(i+1)]\} \tag{2.65}$$

and define a scalar sequence H^i (Hamiltonian)

$$H^i = L^i(x(i),u(i)) + \lambda^T(i+1)f^i(x(i),u(i)) \tag{2.66}$$

$$i = 0,\ldots,N-1 \;.$$

Substituting Equation 2.66 into 2.65 we have

$$\bar{J} = \emptyset(x(N)) - \lambda^T(N)x(N) + \sum_{i=1}^{N-1} [H^i - \lambda^T(i)x(i)] + H^o \ . \tag{2.67}$$

Now consider the differential change in $\bar{J}$ due to differential changes in u(i).

$$d\bar{J} = [\frac{\partial \emptyset}{\partial x(N)} - \lambda^T(N)]dx(N)$$

$$+ \sum_{i=1}^{N-1} \{[\frac{\partial H^i}{\partial x(i)} - \lambda^T(i)]dx(i) + \frac{\partial H^i}{\partial u(i)} du(i)\}$$

$$+ \frac{\partial H^o}{\partial u(0)} du(0) + \frac{\partial H^o}{\partial x(0)} dx(0) \quad . \tag{2.68}$$

One wants to find conditions on x, u and λ such that the standard first order optimality condition $d\bar{J} = 0$ is satisfied. Choose the adjoint multipliers such that:

$$\lambda^T(i) = \frac{\partial L^i}{\partial x(i)} + \lambda^T(i+1) \frac{\partial f^i}{\partial x(i)} \qquad i=0,1,\ldots,N-1 \tag{2.69}$$

$$\lambda^T(N) = \frac{\partial \emptyset}{\partial x(N)} \tag{2.70}$$

one may specify the necessary optimality condition

$$\frac{\partial H^i}{\partial u(i)} = 0 \qquad i=0,1,2,\ldots,N-1 \quad . \tag{2.71}$$

In summary, to find a control-vector sequence u(i), i=0(1)N that produces a stationary value of J, we must solve the two point boundary value problem defined by Equations 2.63, 2.64, 2.69, 2.70 and the optimality conditions Equation 2.71.

Most gradient methods start with solutions that satisfy neither the optimality conditions Equation 2.71 nor the boundary conditions. The algorithms then generate iterative solutions which improve the control trajectory at each iteration.

Given a control trajectory $u^o(i)$ $i = 0,1,\ldots,N-1$, the gradient procedure goes as follows:

1. Integrate the system Equation 2.63 forward in time using $u^o(i)$.

2. At the terminal time evaluate $\lambda(N)$ and using Equation 2.69 and 2.70 integrate the adjoint trajectories backward in time.

3. Using the calculated $x(i)$, $\lambda(i)$ $i = 0,\ldots,N-1$ calculate the Hamiltonian 2.66 and its gradient with respect to $u(i)$.

4. Find the direction of search to minimize the Hamiltonian by using gradient, conjugate gradient or Davidon's method. Also make a one-dimensional search to determine the scale factor in the search direction.

$$u^{k+1}(i) = u^k(i) + \alpha_k p_k \qquad p_k = \text{direction of search} \quad .$$

5. Return to step 1 with a new $u(i)$ trajectory after

possible modifications to the penalty constants.

Another important approach to solving discrete control problems with inequality constraints

$$g^i(x(i),u(i)) \leq 0 \qquad i = 0,1,\ldots,N-1 \qquad (2.72)$$

is to consider them as large time-staged nonlinear programming problems. The system Equations 2.63 form equality constraints and 2.70 form inequality constraints where 2.62 is to be minimized (65).

One may use some penalty function formulation to reduce the constrained problem to an unconstrained one, and sequential unconstrained optimization techniques to solve the problem (22, 23).

A modification of the gradient method could also be utilized with inequality constraints by using SUMT techniques with the original objective function 2.62.

A difficulty encountered here is the size of the nonlinear programming problem if the time horizon is large.

2. Discrete control growth model with inequality constraints

A second approach to a discrete control problem is concerned with the optimal growth and valuation in multisectoral economies in which the technology is of the discrete, activity-analysis, type. This model is due to Bruno (9) and is believed to have a considerable degree of realism and usefulness in the field of development planning. From the computational viewpoint this analysis allows one to solve a large time-staged programming problem in terms of small subproblems. The link between the time periods is provided by system differential equations and adjoint differential equations. Bruno's (9) main concern in the analysis was to give full asymptotic characterization of the optimal time paths, the price behavior, and the nature of choice of alternative activities for infinite horizon models. My interest is to analyze the computational procedures for the model, given a finite time horizon, and to check the feasibility of the neighboring extremal algorithms applied to the model.

Consider the prototype of the general model as a simple fixed proportion two-sector model. An economy produces two goods, a consumption good C and a depreciable capital good I (I = gross investment), with an exponential depreciation rate μ. Each sector uses, as fixed proportion inputs, both capital and a primary factor of production, labor L, which is assumed to grow at an exponentially fixed rate n.

The production technology is assumed to be given by a coefficient matrix A.

$$A = \begin{pmatrix} a_o & a_{o1} \\ a_1 & a_{11} \end{pmatrix} .$$

Introduction of the following notation gives:

$c(t) = C(t)/L(t)$ = consumption per capita,

$z(t) = I(t)/L(t)$ = gross investment per capita,

$z(t) = \dot{k}(t) + \lambda k(t) \qquad \dot{k} = dk/dt$, and

$k(t) = \frac{K(t)}{L(t)}$ = capital-labor ratio,

$\lambda = n + \mu$ = gross rate of growth.

One can then formulate the following optimal control problem:

$$\text{maximize } J = \int_{t_o}^{t_f} c(t)\, e^{-\delta t}\, dt - \frac{PC}{2}\,[k(t_o) - k_{t_f}]^2 \tag{2.73}$$

where δ = time rate of discount and PC is a positive penalty constant and the following constraints.

(labor constraint) $a_o c + a_{o1} z \leq 1$ (2.74)

(capital constraint) $a_1 c + a_{11} z \leq k$ (2.75)

(nonnegative consumption $-c \leq 0$ (2.76)

(nonnegative investment) $-z \leq 0$ (2.77)

(and the differential equation) $\dot{k}(t) = -\lambda k(t) + z(t)$ (2.78)

(boundary conditions) $k(t_o) = k_o k(t_f) = k_{t_f}$. (2.79)

The necessary conditions for the solution to this problem can be derived from either the calculus of variations or Pontryagin's Maximum Principle (58). One introduces the Hamiltonian form,

$$H(c,k,z,t,\pi) = e^{-\delta t}c(t) + e^{-\delta t}\pi(t)[z(t) - \lambda k(t)] \quad . \tag{2.80}$$

H can be interpreted as the net national product per capita where net investment is valued at the demand price for capital $\pi(t)$. All prices are in consumption units.

Applying theorem 23 (58, p. 29) and the related analysis, we conclude that if a program $[c(t),z(t),k(t); t_o \leq t \leq t_f]$ is optimal, then there exists a continuous function $\pi(t)$ such that

$$\dot{\pi}(t) = (\lambda + \delta)\pi - s \quad . \tag{2.81}$$

This is seen from writing inequalities 2.74 and 2.75 as equalities

$$a_o c + a_{01} z + \varepsilon_o - 1 = 0 \quad , \tag{2.82}$$

$$a_1 c + a_{11} z + \varepsilon - k = 0 \tag{2.83}$$

where ε_o and ε are slack functions. Then from theorem 23 we have,

$$\frac{d}{dt}(e^{-\delta t}\pi(t)) = -\frac{\partial H}{\partial k} + e^{-\delta t} w \frac{\partial}{\partial k}[a_o c + a_{o1} z + \varepsilon_o - 1]$$

$$+ s e^{-\delta t} \frac{\partial}{\partial k}[a_1 c + a_{11} z + \varepsilon - k] \quad . \tag{2.84}$$

Equation 2.84 reduces to

$$\dot{\pi}e^{-\delta t} - \delta e^{-\delta t}\pi = -(-\lambda\pi)e^{-\delta t} + s(-1)e^{-\delta t} \quad , \qquad (2.85)$$

and

$$\dot{\pi} = (\lambda + \delta)\pi - s \quad .$$

At each moment of time, gross national product,

$$GNP = He^{\delta t} + \pi\lambda k = c + \pi z \quad , \qquad (2.86)$$

is maximized subject to inequalities 2.74 through 2.77. This is equivalent to solving a linear programming problem at each moment in time. Its dual is:

$$\text{minimize } De^{\delta t} = w + sk \qquad (2.87)$$

where D is discounted gross national income, subject to the constraints;

$$a_o w + a_1 s \geq 1 \qquad (2.88)$$

$$a_{o1} w + a_{11} s \geq \pi \qquad (2.89)$$

$$w \geq 0 \qquad (2.90)$$

$$s \geq 0 \quad . \qquad (2.91)$$

w has the interpretation of the real wage rate and s that of gross rental price of capital. In the former notation k(t) is the state variable, c(t) and z(t) are the control variables, $\pi(t)e^{-\delta t}$ is the auxiliary variable and w and s are the Lagrange multiplier functions. In addition $\pi(t)$

must satisfy the conditions due to the terminal constraint on k(t), namely

$$e^{-\delta t_f}\pi(t_f) = PC[k(t_f) - k_f] \quad . \qquad (2.92)$$

Writing the inequality constraints in equality form one has:

the Production Equations

$a_o c + a_{o1} z + \varepsilon_o = 1$

$a_1 c + a_{11} z + \varepsilon = k$

the Price Equations

$a_o w + a_1 s - P_o = 1$

$a_{o1} w + a_{11} s - P = \pi$

for all t in the interval $t_o \leq t \leq t_f$. The nonnegative slack variables have the following economic interpretation:

ε_o = rate of unemployment of labor

ε = excess capacity per unit of labor

P_o = difference between the supply price and demand price of consumption

P = difference between supply price and demand price of capital.

From linear programming theory we know that we must have

$$w\varepsilon_o = s\varepsilon = P_oC = Pz = 0 \quad . \qquad (2.93)$$

Now within the framework of the model as given one can consider the computational procedure. The matrix A is given as is also t_o and t_f, $k(t_o)$, $k(t_f)$, λ and δ. With $k(t_o)$ given one estimates $\pi(t_o)$ and then solves the primal linear program Equations 2.74 through 2.77 and its dual. This then with objective function 2.86 gives values for $c(t_o)$, $z(t_o)$, $w(t_o)$ and $s(t_o)$. The differential Equations 2.78 $\dot{k} = z-\lambda k$ and Equation 2.81 $\dot{\pi} = (\lambda + \delta)\pi - s$ allow one to step up the time interval for $k(t)$ and $\pi(t)$ to $k(t_1)$ and $\pi(t_1)$. The linear program and its dual are again solved generating $c(t_1)$, $z(t_1)$, $w(t_1)$ and $s(t_1)$ and the iteration continues until $k(t_f)$ is computed. If $k(t_f)$ agrees with k_f then the optimal time paths have been computed, if not then $\pi(t_o)$ must be modified and the process repeated from $t = t_o$ to $t = t_f$. Interpolation can be used to assist in the selection of a proper $\pi(t_o)$. This computational procedure could be extended to models with more sectors. But as the numbers of adjoint variables increase, the problem of selecting proper initial values for these variables becomes increasingly difficult. This procedure is sometimes called the neighboring extremal algorithm (12) and a statement of the method is as follows:

1. Approximate initial values of the control problem adjoint variables at the initial time.

2. Integrate the state and adjoint differential equations forward in time, at the same time make an optimal choice of the control variables using the current values of the state and adjoint variables and observe how far the state variables at terminal time miss the boundary conditions.

3. Using this observation modify the approximation of the initial adjoint variables unless sufficient accuracy has been obtained and go back to step (2) until a convergence criterion has been met.

III. ECONOMIC APPLICATIONS AND NUMERICAL SOLUTIONS

A. Introduction to the Control Problem Computation

In this report, all automatic computations were performed on the IBM 360/40 digital computer using Fortran IV language and double precision arithmetic with accuracy of approximately sixteen decimal digits. All integrations were performed using fourth order numerical integration methods. Fixed stepsize was used in the Runge-Kutta procedure. The interval of integration was divided up into 100 equal subdivisions. Each one-dimensional minimization required in a solution reported here was based upon a cubic polynomial approximation to the contour of the functional along the direction of search. After a satisfactory approximation was made, the positive value corresponding to a minimum of the polynomial was chosen as the optimum search-direction stepsize. This procedure, described earlier, has been used extensively in finite dimensional problems and proved satisfactory here for control problems (26).

At this point I would like to comment about using penalty functions to handle terminal constraints on the state variables. The penalty function approach is an alteration of the form of the optimal control problem itself, rather than a modification of the numerical technique used to solve it. The constrained problem is approximated by one or more unconstrained problems by adding to the objective functional a positive measure of the constraint violation.

The penalty function approach attempts to force those controls producing large constraint violations to lie on contours of higher objective functional values in the control space than those producing smaller constraint violations. The choice of the values of the penalty constants influences the objective functional throughout the entire control space. I have found that for a typical control problem, the effect of the penalty term is extremely difficult if not impossible to determine without numerical experimentation. Therefore in many cases the choice of the values of the penalty constants is arbitrary and must be chosen on the basis of numerical trials.

Some of the difficulties involved in using penalty functions can be avoided by replacing a single solution attempt by a sequence of solutions involving increased weighting of the constraint violation. Each new subproblem is started with the control computed from the previous subproblem. This problem has been studied extensively as mentioned for finite dimensional optimization procedures by Fiacco and McCormick (22, 23, 24). The choice, however, of the penalty constants for each subproblem must still be made arbitrarily at first and modified on the basis of numerical experience with each subproblem.

Either the fixed or the increasing sequence of penalty constants were used to solve the problems that follow. After initial trial and error with penalty constants for each problem, adequate penalty constants were determined to handle state terminal constraints.

A nonlinear control problem with quadratic objective functional was given by Willoughby (71). This problem was used to check out the computer codes for both the conjugate gradient and the Davidon algorithms and the results are included here to illustrate the convergence of the methods. A statement of this test problem, T-1, with the penalty function included follows:

$$\text{minimize } J = 1/2 \int_0^5 (x_1^2 + x_2^2 + u^2)dt + \frac{p}{2}(x_2(5) - x_1(5) - 1.0)^2 \quad (3.1)$$

subject to:

$$\dot{x}_1 = x_2 \qquad x_1(0) = 1 \quad ,$$

$$\dot{x}_2 = -x_1 + (1 - x_1^2)x_2 + u \quad , \quad x_2(0) = 0 \quad ,$$

$$\Omega(x(5)) = -x_1(5) + x_2(5) - 1 = 0 \quad . \qquad (3.2)$$

The initial control estimate was chosen $u_o(t) = 0$ for all t in the interval [0,5]. For the Davidon algorithm four iterations were performed before restarting with a direction of search chosen in the negative gradient direction. The solution to the above problem is presented in Table 3.1. The penalty constant used was 10.0 and the Runge-Kutta step-size was h = .05. In Table 3.2, the result of a sequence of subproblems is presented with variable penalty constants and initial control function $u_0(t) = 0$. The control for each successive subproblem is generated from the preceding subproblem.

B. One-Sector Neoclassical Growth and Optimal Growth Models

To introduce the computation of the optimal growth model first consider the growth model, with no objective functional, which characterizes economic growth in an aggregate closed economy. Aggregate means that the economy produces a single homogeneous good, the output at time t is Y(t), using two inputs, labor L(t) and capital K(t). The adjective, closed, refers to the point that neither output nor input is imported or exported. All output from the productive process is either consumed or invested. If one represents consumption as C(t) and investment at I(t) then the income identity can be written as

$$Y(t) = C(t) + I(t) \quad , \qquad (3.3)$$

which states that output (Gross National Product) can either be consumed or invested.

Investment is used to increase the stock of capital

Table 3.1. Penalty function solution of problem T-1 using a fixed penalty constant of p = 10.0

Iteration Number	Davidon Method			Conjugate Gradient Method		
	J	Ω(x(5))	(g,g)	J	Ω(x(5))	(g,g)
1.	7.8901	.0605	15.6090	7.8901	.0657	15.6092
2.	2.1788	-.0089	15.6929	2.1749	-.0126	14.5049
3.	2.1532	-.0674	7.2383	2.1561	-.0702	5.3084
4.	1.9949	-.0656	1.7995	2.1447	-.1158	7.9474
5.	1.9820[a]	-.1048	3.2139	2.1069	-.1599	12.4029
6.	1.9327	-.0520	4.2029	2.0890	-.1252	4.6076
7.	1.6746	-.0585	.3276	2.0804	-.0885	4.2695
8.	1.6722	-.0739	.8537	2.0652	-.0357	10.3254
9.	1.6712	-.0571	.0700	2.0182	-.0388	29.6590
10.	1.6707[a]	-.0562	.0020	1.9164	-.0280	28.7799
11.	1.6701	-.0551	$.2\times10^{-4}$	1.8810	-.0391	12.0685
12.	1.6701	-.0551	$.2\times10^{-5}$	1.8649	-.0916	19.0845
13.	1.6701	-.0551	$.2\times10^{-5}$	1.7856	-.1988	115.4830
14.				1.7404	-.1552	51.0380
15.				1.7277	-.1172	20.2462
16.				1.7227	-.0889	9.8346

[a]Result of a negative gradient direction of search.

Table 3.2. Results of Davidon and conjugate gradient algorithm applied to problem T-1 with variable penalty constants, initial estimate $u_o = 0.0$

Subproblem Number	Penalty Constant	$J - \frac{p}{2}\,\Omega^2(x(5))$	$\Omega(x(5))$	Number of Steps Taken
		Davidon Method		
1	10.0	1.645	-.07	8
2	50.0	1.645	-.07	3
3	100.0	1.6863	-.005	3
		Conjugate Gradient Method		
1	10.0	2.0590	-.035	8
2	50.0	1.9032	-.0030	3
3	100.0	1.7059	-.0009	3
4	200.0	1.7006	-.0002	2

and to replace depreciated capital. Letting K(t) be the stock of capital at time t and assuming that the stock of capital depreciates at a rate δ, then gross investment identity states that:

$$I(t) = \dot{K}(t) + \delta K(t) \quad . \tag{3.4}$$

Capital accumulation is that part of investment not used to replace depreciated capital.

Output is determined by an aggregative production function which summarizes the technically efficient possibilities for production of output from capital and labor:

$$Y(t) = F(K(t),L(t)) \tag{3.5}$$

$$F_K > 0 \ , \quad F_L > 0 \ , \quad F_{KK} < 0 \ , \quad F_{LL} < 0$$

$$\lim_{K\to 0} F_K = \infty \ , \quad \lim_{K\to\infty} F_K = 0$$

$$\lim_{L\to 0} F_L = \infty \ , \quad \lim_{L\to\infty} F_L = 0 \quad . \tag{3.6}$$

Also if one assumed that the production function exhibits constant returns to scale, then

$$\frac{Y}{L} = F(\frac{K}{L},1) = f(\frac{K}{L}) = f(k) \quad , \tag{3.7}$$

where the function f(.) gives output per worker as a function of capital per worker. Denote per worker quantities by lower case letters:

$$y(t) = Y(t)/L(t) \ , \quad k(t) = K(t)/L(t)$$

$$c(t) = C(t)/L(t) \ , \quad i(t) = I(t)/L(t) \quad ,$$

by Equation 3.4 $f'(k) > 0$, $f''(k) < 0$, $\forall k$

$$\lim_{k\to 0} f'(k) = \infty \ , \quad \lim_{k\to\infty} f'(k) = 0 \quad .$$

The labor force is assumed to grow at the given exponential rate r

$$\dot{L} = rL \quad . \tag{3.8}$$

The income identity, the gross investment identity and the production function can be combined in per worker terms to form the fundamental differential equation of neoclassical economic growth,

$$f(k(t)) = c(t) + \lambda k(t) + \dot{k}(t) \quad , \tag{3.9}$$

where $\lambda = r + \delta$. This differential equation states that output per worker $f(k)$ is allocated among three uses:

1. Consumption per worker $c(t)$,
2. Maintenance of the level of capital per worker $\lambda k(t)$,
3. Net increase in the level of capital per worker $\dot{k}(t)$.

Two values $\hat{k}$ and $\tilde{k}$ designate levels of capital per worker at which $c + \dot{k}$ is a maximum and zero respectively.

$$f(\hat{k}) - \lambda\hat{k} \geq f(k) - \lambda k \qquad \forall k > 0$$

$$f(\tilde{k}) - \lambda\tilde{k} = 0 \quad . \tag{3.10}$$

Under the assumption given $\hat{k}$ and $\tilde{k}$ exist and are unique (36).

$$f'(\hat{k}) = \lambda = \delta + r \quad . \tag{3.11}$$

The maximized level of consumption per worker $\hat{c}$ that can be maintained forever as an equilibrium level at $\hat{k}$ is given by,

$$\hat{c} = f(\hat{k}) - \lambda\hat{k} \quad , \tag{3.12}$$

where $\hat{c}$ is called golden-rule level of consumption per worker. Condition 3.11 is called the golden rule of accumulation. $\hat{k}$ is an equilibrium but not a stable equilibrium. Deviations to the right of $\hat{k}$ are eliminated but to the left are not (36).

The problem of optimal economic growth is a dynamic control problem. In the one sector problem there is one state variable k(t), capital per worker and the equation of motion is the fundamental differential equation of neoclassical economic growth.

$$\dot{k} = f(k) - \lambda k(t) - c(t)$$

$$k(t_o) = k_o \quad k(t_f) = k_f \quad . \tag{3.13}$$

From the viewpoint of a central planner who has authority over the entire economy, the control variable is consumption per worker. The problem then is that of choosing a time path for consumption per worker over the planning horizon:

$$\{c(t) = c(t) \mid t_o \leq t \leq t_f\} \tag{3.14}$$

where t_o, t_f, f(.), λ, k_o, k_f, are assumed given. Any time path satisfying the differential Equation 3.13 and the boundary condition for which,

$$0 \leq c(t) \leq f(k(t)) \quad \forall \quad t \varepsilon [t_o, t_f] \quad ,$$

is feasible and the problem facing the central planner is that of choosing a feasible trajectory for consumption per worker that is optimal in achieving some economic objective.

The economic objective of the central planner is assumed to be based on standards of living as measured by consumption per worker. In particular it is assumed that the planner has a utility function $u(c(t))$, giving utility at any time as a function of consumption per worker or a disutility function measuring the squared deviation from some desired time path of consumption. It is assumed that utilities at different times are independent and that utilities at different times can be added, after they have been suitably discounted to allow for the fact the near future generations are politically more important than far future generations. The rate of discount, ρ, assumed constant and nonnegative, is the marginal rate of transformation between present and future utility.

The problem of neoclassical optimal growth for an aggregate closed economy with a finite time horizon and positive discount rate and the assumptions on the production function previously mentioned is that of choosing a time path for consumption per worker, $c(t)$, such that the following equations are satisfied.

$$\text{maximize } J = \int_{t_o}^{t_f} e^{-\rho t}\, u(c(t))dt \quad , \qquad (3.15)$$

$$\dot{k} = f(k) - \lambda k - c \quad , \tag{3.16}$$

$$k(t_o) = k_o \quad , \quad k(t_f) = k_f \quad ,$$

$$0 \leq c(t) \leq f(k) \quad \forall \quad t \epsilon [t_o, t_f] \quad . \tag{3.17}$$

c(t) piecewise continuous,

$\lambda = r + \delta$.

The solution to this problem is an optimal path for consumption per worker $c^*(t)$ and an optimal path for capital per worker $k^*(t)$ for all $t \epsilon [t_o, t_f]$. The solution depends upon two functions f(.) and u(.), on the nonnegative parameters, 1. rate of discount ρ, 2. depreciation rate plus growth rate of labor, $\lambda = \delta + r$, 3. initial stock of capital, 4. final stock of capital.

The Hamiltonian for the problem can be written,

$$H(k, c, \pi, t) = e^{-\rho t}[u(c) + \pi(f(k) - \lambda k - c)] \tag{3.18}$$

where the adjoint variable is $\pi(t)e^{-\rho t}$.

The term in the brackets is the sum of utility and the adjoint variable multiplied by the net investment per worker, indicating an interpretation of $\pi(t)$ as the inputed value (shadow price) of additional capital per worker, measured in terms of utility. The Hamiltonian is the inputed value discounted to the initial time zero.

As an initial sequence of numerical experiments illustrating the control algorithms applied to solve economic models numerically, one may formulate a model similar to the type studied by Goodwin (30). This particular model has linear production and capital accumulation functions, but the technique of obtaining the numerical solution is in no way restricted to linear cases. These functions were selected only as an initial illustration and will be followed by studies of nonlinear relationships. This model differs from that of Goodwin in that it has a quadratic valuation function of the squared difference between per worker consumption $c(t)$ and a known desired per worker consumption $c^*(t)$ rather than a log function. The function $c^*(t)$ may be a derived function from optimizing on the subunit level or it may arise from the subjective preferences of the planners or possibly a subset of the planners.

Suppose for example that a group within the economic unit, say the businessmen, or a sectoral group want $c^*(t)$ to have a certain time path subject to the dynamic constraints of production and capital accumulation. They, however, would accept as a compromise a path close to their desired path in terms of the minimum of a squared deviation from $c^*(t)$. The objective is to choose $c(t)$ as close to $c^*(t)$ as possible subject to the constraints of the model. The variables are defined as:

$K(t)$ = aggregate quantity of the capital of the economic unit,

$C(t)$ = aggregate consumption of the unit,

$L(t)$ = labor force = $L_o e^{rt}$,

$k(t) = K(t)/L(t)$,

$c(t) = C(t)/L(t)$,

$Y(t)$ = output of the economic unit,

$y(t) = Y(t)/L(t)$,

$[t_o, t_f]$ = planning horizon,

K_o = initial capital stock,

K_f = final capital stock,

B = output-capital ratio,

p = penalty constant.

Problem T-2 can then be formulated as follows:

$$\text{minimize } J = \int_{t_o}^{t_f} (c(t) - c^*(t))^2 dt + \frac{p}{2}(K(t_f) - K_f)^2 \tag{3.19}$$

$$\text{subject to: } \dot{K}(t) = Y(t) - L_o e^{rt} c(t) \ , \tag{3.20}$$

$$Y(t) = BK(t) \ , \tag{3.21}$$

$$K(t_o) = K_o \ , \quad K(t_f) = K_f \ . \tag{3.22}$$

The T-2 optimal solution may be computed directly from the above formulation or computed after the problem has been stated in per worker terms. For a representative parameter specification let $r = .01$, $B = .25$ and $L_o = 10.00$. Let the desired control $c^*(t)$ be a given as a subsistence level plus a linear time trend, $c^*(t) = 9.0 + .5t$, and $t_o = 0.0$, $t_f = 10.0$. If one allows for a 5 percent per year rate of growth of output from the economic unit, then $Y(10) = 165.0$, where $Y(0) = 100.0$.

This class of problems, linear dynamics, nonautonomous with quadratic objective functional and state variable terminal constraints, represents one of the easier types of control problems to solve, yet it is important in my analysis since certain types of two and higher sector models, as will be considered later can be reduced to a problem like problem T-2 but with a time varying output-capital ratio. Both the conjugate gradient and the Davidon algorithms were used to solve the problem T-2. In terms of the output variable Y(t) and the adjoint variable $\pi(t)$, the necessary conditions are:

$$\dot{Y}(t) = B(Y(t) - c(t)\,L_o\,e^{rt}) \quad , \quad Y(0) = 100.0 \tag{3.23}$$

$$\dot{\pi}(t) = -B\pi(t),\ \pi(10) = p(Y(10) - 165.0) \tag{3.24}$$

$$g = H_c = 2(c(t) - c^*(t)) - \pi(t)\,L_o\,Be^{rt} = 0 \ , \tag{3.25}$$

where

$$H = (c-c^*)^2 + \pi(t)(Y(t) - L_o\,c(t)e^{rt})B \ . \tag{3.26}$$

The penalty constant used was 3.0 and for both algorithms the initial control used was $c_o(t) = 9.0$. The stopping rule was a value of (g,g) less than 1.0×10^{-4}. Values of the functional J, (g,g) and the number of forward and backward integrations per iteration are summarized in Table 3.3. The conjugate gradient method with this and other experiments was much more sensitive to the α-search direction parameter. It required 44 integrations of the state and adjoint differential equations. Most of these were required to determine the search direction parameter.

The Davidon algorithm was much less sensitive to the search direction parameter. It converged after three steps and 13 integrations of the differential equations. Both methods gave essentially the same results for the trajectories for problem T-2. Results for various time points are given in Table 3.4. The stepsize for the Runge-Kutta integration was h = .1. An approximation of the computation time for the Davidon Algorithm was 18 seconds per iteration. This includes CPU time and printing time. The time per iteration varies depending on how many linear searches must be completed in the iteration to compute an optimal search parameter.

The trajectories for the time horizon of 20 and 30 years respectively for problem T-2 are listed in Tables 3.5 and 3.6.

Table 3.3. Convergence results for problem T-2

Iteration Number	Numbers of Integrations	J	(g,g)	Y(10.0) - 165.0
		Davidon Method		
1.	4	80.5861	263.433	-.0591
2.	7	14.7394	28.205	-.0988
3.	2	14.7370	.000001	-.0584
		Conjugate Gradient Method		
1.	4	80.5861	263.432	-.0591
2.	8	16.2077	16570.00	-.0103
3.	3	14.7370	.0110	-.0592
4.	9	14.7370	.00016	-.0544
5.	5	14.7370	.0246	-.0572
6.	2	14.7370	.0557	-.0566
7.	3	14.7370	.0003	-.0582
8.	10	14.7370	.000003	-.0584

One notes that the savings rate $S(t)/Y(t)$ for the different time horizon differs. In the 10 and 20 year plans the rate decreases monotonically, while in the 30 year plan it drops to approximately .15 in the year 10 and remains there until year 20 and then builds up to satisfy the terminal capital constraint.

Table 3.4. Optimal trajectories for problem T-2 with time horizon [0,10]

t	Y(t)	c(t)	C(t)	S(t)	S(t)/Y(t)
0.0	100.00000	6.32990	63.29900	36.70100	0.36701
1.2	110.53299	7.59810	76.89827	33.63472	0.30430
2.0	117.08499	8.34780	85.16438	31.92061	0.27263
2.8	123.31299	9.03640	92.92993	30.38306	0.24639
3.6	129.24599	9.67470	100.29329	28.95270	0.22401
4.4	134.89699	10.27130	107.33322	27.56377	0.20433
5.2	140.26900	10.83350	114.11751	26.15149	0.18644
6.0	145.35199	11.36750	120.70422	24.6477	0.16957
6.8	150.11800	11.87800	127.13794	22.98006	0.15308
7.6	154.52800	12.36920	133.45900	21.06900	0.13634
8.4	158.52399	12.84450	139.70041	18.82358	0.11874
9.2	162.02800	13.30660	145.88889	16.13911	0.09961
10.0	164.94199	13.75700	152.03839	12.90359	0.07823

Table 3.5. Optimal trajectories for problem T-2 with time horizon [0,20]

t	Y(t)	c(t)	C(t)	S(t)	S(t)/Y(t)
0.0	100.00000	6.0300	60.29999	39.70001	0.39700
2.4	122.17999	8.5300	87.37193	34.80806	0.28489
4.0	135.73000	9.86300	102.65514	33.07486	0.24368
5.6	148.75000	11.02500	116.60017	32.14983	0.21613
7.2	161.51999	12.07000	129.71091	31.80908	0.19694
8.8	174.23999	13.03900	142.38425	31.85574	0.18283
10.4	187.39999	13.95000	154.78920	32.61079	0.17402
12.0	200.03999	14.83000	167.20775	32.83224	0.16413
13.6	213.31999	15.68500	179.70010	33.61989	0.15760
15.2	226.95999	16.51999	192.31906	34.64093	0.15263
16.8	241.04999	17.34999	205.23940	35.81059	0.14856
18.4	255.67999	18.16199	218.30997	37.37003	0.14616
20.0	271.00000	18.97299	231.73671	39.26329	0.14488

Table 3.6. Optimal trajectories for problem T-2 with time horizon [0,30]

t	Y(t)	c(t)	C(t)	S(t)	S(t)/Y(t)
0.0	100.00000	6.02600	60.25998	39.74002	0.39740
3.6	132.45999	9.54600	98.95912	33.50087	0.25291
6.0	152.06999	11.29500	119.93439	32.13560	0.21132
8.4	171.26999	12.80300	139.24905	32.02094	0.18696
10.8	190.64999	14.17700	157.93857	32.71143	0.17158
13.2	210.65999	15.47400	176.57512	34.08487	0.16180
15.6	231.68999	16.72899	195.53278	36.15721	0.15606
18.0	254.25000	17.95999	215.02007	39.22993	0.15430
20.4	279.08984	19.17699	235.16705	43.92279	0.15738
22.8	307.51978	20.38699	256.07788	51.44189	0.16728
25.2	341.82178	21.59200	277.80176	64.02002	0.18729
27.6	386.14380	22.79500	300.40308	85.74072	0.22204
30.0	447.98584	23.99699	323.92529	124.06055	0.27693

Over the ten year time horizon one can obtain a regression of c(t) against Y(t) to determine the control variable as a function of the state variable. A linear fit of these data gives the relationship c(t) = -4.65605 + .110624 Y(t), with a coefficient of multiple determination, R^2 value, of .99927 and a residual variance of .0037328.

Letting $t_{i+1} - t_i = .4$, a lagged relationship between Y_{t-1} and c_t for the interval [0,10] is given by,

$$c_t = -3.812296 + .1065648\ Y_{t-1}$$

with an R^2 value of .999425 and a residual variance of

.00236.

For the twenty year horizon a linear relationship of the data is given by the equation $c(t) = -.241 + .0736\ Y(t)$ relating consumption per worker and aggregate output with an R^2 value of .9865. For the thirty year horizon the above data relationship appears quadratic $c(t) = -4.1781 + .1201\ Y(t) - .000129Y^2(t)$ with an R^2 value of .99879. This would imply that, using model T-2, a linear rule giving the consumption per worker as a function of the output would only be valid within a 20 year horizon.

It appears that most practical planning situations would be within a small time horizon, since one may not be able to obtain deterministic relationships over a long horizon. For the T-2 problem c(t) approaches c*(t) in approximately 15 years, hence the transient terms are necessary in this realistic optimal short range planning.

Control problem T-2 can be resolved analytically. From the necessary conditions one determines c(t) and Y(t) as,

$$c(t) = 9.0 + .5t + (1.25)(A)e^{-.24t} \qquad (3.27)$$

$$Y(t) = Be^{.25t} + 5.208te^{.01t} + 115.45e^{.01t} + (6.51)(A)e^{-.23t} \qquad (3.28)$$

Where A and B are constants to be determined by the boundary conditions. For the time horizon [0,10] and the given parameters of the model, the constants have the following values, A = -2.13696 and B = -1.5389. The value of J = 14.7429 and $J - \frac{3.0}{2.0}(y(10.0 - 165.)^2 = 14.7375$. This compared with the computed value of 14.7370. This problem T-2 was a good test of the computer code and also it indicated the ease with which different parameter settings can be made and the solution obtained by the numerical algorithms used. This type of problem T-2 will be utilized again in a suboptimization procedure of a two-sector model. The suboptimization procedure will involve a time varying output-capital ratio and will be discussed in Section C.

As a second sequence of numerical experiments, consider the following model. This model was studied by Chakravarty (16) and he considered a nonlinear welfare function and a nonlinear production function. He showed that for a production function of the form, $Y = aK^B$ where Y is the output, K is the capital stock, and a and B are parameters that if B = 1/2 it was possible to obtain a closed form solution for the time path of capital stock. The B = 1/2 case was the only nonlinear problem he discussed since he implied that it was not possible to obtain closed form solutions for any other cases. This model does not express the variables in per worker terms. A modification of this model where per worker variables are considered will be treated later.

The control problem T-3 is formulated as follows:

$$\text{maximize } J = \int_{t_o}^{t_f} e^{-\rho t} \frac{1}{1-n}(C(t))^{1-n} dt \qquad (3.29)$$

subject to

$$\dot{K}(t) = e^{zt}\gamma K(t)^B (L_o e^{rt})^{1-B} - C(t) - \delta K(t) \quad , \qquad (3.30$$

$$K(0) = K_o \text{ and } K(t_f) = K_f \quad . \qquad (3.31)$$

Where the variables are:

J = an index of performance,

ρ = time rate of welfare discount,

$C(t)$ = consumption at time t,

n = elasticity of marginal utility with respect to consumption,

$\dot{K}(t)$ = capital accumulation,

$K(t)$ = stock of capital,

δ = rate of capital depreciation,

$K(0) = K_o$ is the initial stock of capital,

$K(t_f) = K_f$ is the terminal stock of capital,

z = rate of neutral technical progress,

γ = efficiency parameter,

B = elasticity of output with respect to capital,

L_o = initial labor force,

r = rate of growth of the labor force,

$[0,t_f]$ = fixed time horizon.

The form of the production function used is

$$Y(t) = F(K(t)) = e^{zt}\gamma(K(t))^{B}(L_{o}e^{rt})^{1-B} \quad ,$$

$$= ae^{gt}(K(t))^{B} \quad , \qquad (3.32)$$

where

$$a = \gamma L_{o}^{1-B} \quad \text{and} \quad g = r(1-B) + z \quad .$$

The utility function

$$U(C(t)) = \frac{1}{1-n}(C(t))^{1-n} \quad \text{where} \quad n \geq 0 \quad \text{and} \quad n \neq 1 \qquad (3.33)$$

has the following properties:

$$U'(C(t)) \geq 0 \qquad C \geq 0$$

$$U''(C(t)) \leq 0 \qquad C \geq 0$$

$$\lim_{n \to 0} U(C(t)) = C(t) \quad .$$

An attempt has been made to gain insight into how nonlinear specification of these functions affects the time paths of the optimal solution trajectories. Penalty functions are used to handle terminal constraints on the state variables. The parameter values for the model are given in Table 3.7.

Selected values of the optimal trajectories for problem T-3 are given in Table 3.8.

The value of the functional was 98.182 and a fixed

Table 3.7. Parameter values for Model T-3

δ = .05	ρ = .03	n = .9	r = .025
K_o = 15.0	z = .01	Y_o = 4.27	L_o = 15.0
γ = .285	$Y_f = F(K_f) = 7.04$[a]		

Complementary values of a and B

B	a
.60	.8419
.75	.5601
.80	.4900
1.00	.2850

[a]This allows a 5% growth per year in output over the planning horizon

Table 3.8. Optimal trajectories for problem T-3 with B = .6, a = .8419 and time horizon of 10

Time t	Output Y(t)	Capital K(t)	Consumption C(t)	Adjoint Variable $\pi(t)$	Savings Rate
0.0	4.275	15.000	2.255	.479	.472
.4	4.395	15.503	2.369	.457	.461
1.2	4.636	16.497	2.575	.415	.445
2.0	4.878	17.485	2.769	.377	.432
2.8	5.122	18.470	2.973	.344	.419
3.6	5.367	19.440	3.200	.314	.403
4.4	5.610	20.374	3.455	.286	.384
5.2	5.845	21.248	3.737	.260	.361
6.0	6.071	22.039	4.043	.238	.334
6.8	6.283	22.722	4.369	.217	.304
7.6	6.477	23.277	4.709	.197	.273
8.4	6.651	23.683	5.060	.179	.239
9.2	6.798	23.921	5.419	.163	.202
10.0	6.917	23.970	5.783	.148	.164
J = 98.182					

penalty constant of 5.0 was used with an initial control variable of $C_o(t) = 1.0$. In all the numerical experimentation the Davidon algorithm was much less sensitive to both the initial control estimate and the search direction parameter. In every case it converged with less iterations than the conjugate gradient method. The restart feature of the Davidon method proved to be an asset rather than a practical necessity. Restarting the search direction in a negative gradient direction every 4 to 6 iterations proved completely adequate in my computational experience.

The shadow price of additional capital measured in terms of utility is seen to start out at .479 and decreased to .148 as the terminal constraint on capital is satisfied. The savings rate decreases from .472 to .164 over the 10 year horizon.

The elasticity of output with respect to capital, B, is now varied while holding all other parameters constant. In varying B the parameter "a" is chosen in a complementary manner to maintain a constant initial level of output Y(t) with the different specifications of the production function. The optimal saving rates at various time points are computed for different B values and presented in Table 3.9.

The behavior of the savings rate agrees with the expectation that when attempting to hit a certain target rate of growth of output (in the example of problem T-3 5% per

Table 3.9. Problem T-3 optimal savings rate with various values of B and a and planning horizon of 10 years

t time	B=.6 a=.8419	B=.75 a=.5609	B=.8 a=.4900	B=1.0 a=.285
0.0	.472	.548	.609	.739
.4	.461	.541	.578	.664
1.2	.445	.518	.553	.637
2.0	.432	.494	.531	.626
2.8	.419	.467	.513	.589
3.6	.404	.453	.477	.523
4.4	.384	.424	.435	.482
5.2	.361	.389	.394	.421
6.0	.334	.344	.344	.329
6.8	.305	.287	.279	.220
7.6	.273	.217	.197	.100
8.4	.239	.135	.097	.033
9.2	.202	.038	.002	.002
10.0	.164	.000	.000	.000

year), an economic unit with more productive capital should save more in the earlier years of the planning horizon. This example also illustrates the need for obtaining good estimates in the production function parameters as the optimal trajectories change with respect to different values of the parameter B.

Table 3.10 shows changes in the savings rate under variations in n. All other parameters are as given in Table 3.7 with B=.6 and a=.8419. The optimal savings rate for various

Table 3.10. Savings rate and the functional value for different values of n for problem T-3 (B=.6 and a=.8419)

t	n=.2	n=.6	n=.8	n=.9
0.0	.543	.573	.493	.472
1.2	.482	.540	.474	.444
2.4	.434	.499	.445	.426
3.6	.387	.456	.418	.404
4.8	.347	.406	.384	.373
6.0	.309	.337	.338	.334
7.2	.268	.252	.280	.289
8.4	.233	.153	.213	.239
9.6	.192	.035	.134	.184
10.0	.178	.000	.105	.164
	J=30.78	J=36.27	J=55.84	J=98.18

time points are summarized in Table 3.10.

Changes in the values of n appear to have relatively less effect on the savings rate than do changes in B.

Using the notation of the control problem defined in Equations 2.1 and 2.2, consider now the change in the Hamiltonian over time. Since in general H is a function of x, u, λ and t, one may compute $\frac{dH}{dt}$ as follows, where in a one sector model all functions are scalar functions.

$$\frac{dH}{dt} = \frac{\partial H}{\partial x}\dot{x} + \frac{\partial H}{\partial u}\dot{u} + \dot{\lambda}\frac{\partial H}{\partial \lambda} + \frac{\partial H}{\partial t}$$

$$\frac{dH}{dt} = [\frac{\partial H}{\partial x} + \dot{\lambda}]f(x,u,t) + \frac{\partial H}{\partial u}\dot{u} + \frac{\partial H}{\partial t} \quad .$$

Along the optimal trajectory the first term vanishes because of the adjoint differential equations. The second term vanishes because either the partial derivative $\frac{\partial H}{\partial u} = 0$ for an interior solution or $\dot{u} = 0$ for a boundary solution. Thus, along the optimal trajectory $\frac{dH}{dt} = \frac{\partial H}{\partial t}$. If the problem is autonomous in that both L and f show no explicit dependence on time, then $\frac{dH}{dt} = 0$ and along the optimal trajectory the value of the Hamiltonian is constant over time.

Problem T-3 is not an autonomous problem, since the Hamiltonian function depends explicitly on time. To see how the Hamiltonian function behaves for the T-3 problem, for selected time points its value was computed for certain feasible values of the control and state variables and the optimal values. These are presented in Table 3.11 together with (g,g) where,

$$(g,g) = \int_0^{10} \left[\frac{\partial H}{\partial u}^T \frac{\partial H}{\partial u}\right] dt \quad .$$

Next a series of computations with different growth rates on output Y(t) were considered. The results of this experimentation are given for selected time points in Table 3.12 and Table 3.13. The values of the parameters are as given in Table 3.7 with Y_f, final output, computed with 5%, 6%, 7% and 8% growth rate per year.

As seen by Table 3.12 and 3.13, the time paths of consumption and saving rate vary with respect to changes in the

Table 3.11. Values of the Hamiltonian for selected time points for problem T-3 with B=.75 and a=.5609

t	4th iteration	8th iteration	12th iteration
0	12.21	11.39	11.57
1.0	11.87	11.12	11.29
2.0	11.52	10.88	11.02
3.0	11.14	10.63	10.74
4.0	10.76	10.37	10.47
5.0	10.39	10.12	10.20
6.0	10.04	9.89	9.94
7.0	9.70	9.65	9.67
8.0	9.39	9.40	9.40
9.0	9.09	9.16	9.15
10.0	8.81	8.92	8.89
	(g,g)=.368	(g,g)=.042	(g,g)=$.5\text{x}10^{-4}$

Table 3.12. Time paths of consumption where final target output Y_f is computed using different growth rates per year

t	C(t) 5%	C(t) 6%	C(t) 7%	C(t) 8%
0.0	2.25	2.12	1.96	1.62
1.2	2.57	2.37	2.18	1.90
2.0	2.76	2.56	2.34	2.05
3.2	3.08	2.88	2.63	2.27
4.4	3.46	3.19	2.95	2.52
5.6	3.90	3.55	3.26	2.77
7.2	4.55	4.12	3.70	3.11
8.8	5.23	4.79	4.24	3.52
10.0	5.75	5.32	4.71	3.91
	J=98.18	J=97.37	J=96.44	J=94.90

Table 3.13. Time paths of output Y(t) and savings rate s(t) for target final output Y_f determined from different growth rates

t	Y(t) 5%	s(t) 5%	Y(t) 6%	s(t) 6%	Y(t) 7%	s(t) 7%	Y(t) 8%	s(t) 8%
0.0	4.27	.48	4.27	.50	4.27	.54	4.27	.62
1.2	4.64	.45	4.67	.49	4.71	.54	4.77	.60
2.0	4.88	.43	4.94	.47	5.02	.53	5.12	.60
3.2	5.25	.41	5.36	.46	5.50	.52	5.68	.60
4.4	5.61	.38	5.79	.45	5.99	.51	6.27	.59
5.6	5.96	.34	6.23	.43	6.51	.49	6.90	.59
7.2	6.38	.29	6.81	.39	7.23	.49	7.83	.60
8.8	6.72	.22	7.37	.35	7.99	.47	8.85	.60
10.0	6.91	.16	7.77	.31	8.59	.45	9.67	.60

growth rate of the final target output Y_f. The saving rate in Table 3.13 for an 8% per year rate of growth is seen to be almost constant at .60. Certainly if an economic unit can survive on the low time path of consumption as in the 8% per year growth rate their potential for future consumption would increase.

As seen in Table 3.14 the more that capital is needed to attain the various growth rates on the target final output, the larger is the value of the adjoint variable value or the shadow price of capital.

Problem T-3 can be analyzed in per worker terms by making the following changes. The new variables are defined as:

Table 3.14. Time path of adjoint variables with different final target output growth rates

t	π(t) 5%	π(t) 6%	π(t) 7%	π(t) 8%
0	.4781	.5112	.5505	.6251
1.2	.4141	.4431	.4773	.5425
2.0	.3768	.4036	.4352	.4953
3.2	.3276	.3517	.3802	.4343
4.4	.2852	.3073	.3335	.3827
5.6	.2484	.2690	.2933	.3388
7.2	.2065	.2257	.2482	.2899
8.8	.1710	.1895	.2108	.2498
10.0	.1478	.1661	.1869	.2243

$c(t) = C(t)/L(t)$ = consumption per worker,

$k(t) = K(t)/L(t)$ = capital per worker,

$i(t) = I(t)/L(t)$ = investment per worker.

Using the utility function where argument is per worker consumption, the performance functional with the penalty term becomes,

$$\text{maximize } J = \int_{t_o}^{t_f} e^{-\rho t} \frac{1}{1-n}(c(t))^{1-n} dt - \frac{P}{2}(k(t_f) - k_f)^2 \ .$$

Substituting $\dot{K} = \dot{k}L + k\dot{L}$ into the Equation 3.30 and dividing by L, the per worker capital accumulation differential equation is derived.

$$\dot{k} = e^{zt}\gamma k^B - c - (\delta+r)k \tag{3.34}$$

$$0 < B \leq 1$$

$$k(0) = \frac{K(0)}{L(0)} = k_o \quad \text{and} \quad k(t_f) = \frac{K(t_f)}{L(t_f)} = k_f \quad .$$

If the adjoint variable is defined as $\pi = qe^{-\rho t}$, then the Hamiltonian function can be written as,

$$H(k,q,c,t) = e^{-\rho t}[\frac{c^{1-n}}{1-n} + q(e^{zt}\gamma k^B - c - (\delta+r)k)] \quad . \tag{3.35}$$

The adjoint differential equation is:

$$\frac{d}{dt}(qe^{-\rho t}) = - \frac{\partial H}{\partial k} \quad , \tag{3.36}$$

implying that

$$\dot{q} = q[(\delta + r + \rho) - \gamma B e^{zt} k^{B-1}] \quad . \tag{3.37}$$

The first order condition for an interior minimum,

$$\frac{\partial H}{\partial c} = 0 \quad \text{implies that} \quad q = c^{-n} \quad . \tag{3.38}$$

Differentiating 3.38 with respect to time and substituting into 3.37, the two differential equations that the optimal trajectories $\{c(t),k(t)\}$ must satisfy are derived,

$$-n\frac{\dot{c}}{c} = [(\delta+r+\rho) - e^{zt}\gamma Bk^{B-1}] \tag{3.39}$$

$$\dot{k} = \gamma e^{zt}k^B - c - (\delta+r)k \quad .$$

Suppose we allow $z = 0$, assuming no neutral technical progress, $t_f \to \infty$, and temporarily ignore the condition of a given initial stock of capital per worker. Then one possible solution to Equation 3.39 is that for which neither consumption per worker nor capital per worker change over time.

$$\dot{c} = \dot{k} = 0 \quad .$$

In order that consumption per worker be constant it is necessary from Equation 3.39 that $k = k'$, where

$$\gamma B k'^{B-1} = \delta + r + \rho \quad ; \qquad (3.40)$$

and capital per worker will remain at k' if consumption per worker is

$$c' = \gamma k'^{B} - (\delta + r)k' \quad . \qquad (3.41)$$

The equilibrium $k(t) = k'$ and $c(t) = c'$, thus satisfies all the necessary conditions except the initial boundary conditions. This equilibrium $\{k', c'\}$ is the balanced growth path, since along it capital per worker and consumption per worker are constant. Hence total consumption $C(t) = c(t)L(t)$, total capital $K(t) = k(t)L(t)$ and total output $Y(t) = f(k)L(t) = \gamma k^B L(t)$ all grow at the same rate, namely the rate of growth of the labor force. The balanced growth path is called the modified golden rule growth path, since it modifies the golden rule to allow for nonzero discount rate.

$$\lim_{\rho \to 0} k_{,} = \hat{k} \quad . \qquad (3.42)$$

If one assigns the parameter values B = .6, δ = .05, r = .025, γ = .285, ρ = .03 as in Table 3.7 and z = 0.0, then the balanced growth paths may be computed from Equations 3.40 and 3.41 as follows,

$$(.285)(.6)k_{,}^{-.4} = .05 + .025 + .03$$

$$k_{,} = 3.385$$

$$c_{,} = (.285)(3.385)^{.6} - (.075)(3.385)$$

$$c_{,} = .338 \quad .$$

Now consider the optimal path when explicit account is taken of the initial condition on capital per worker and z = 0.

From the differential Equations 3.39

$$\dot{c} = 0 \quad \text{if} \quad \gamma Bk^{B-1} = \delta + r + \rho$$

$$\dot{c} > 0 \quad \text{if} \quad \gamma Bk^{B-1} > \delta + r + \rho$$

$$\dot{c} < 0 \quad \text{if} \quad \gamma Bk^{B-1} < \delta + r + \rho$$

or

$$\dot{c} = 0 \quad \text{if} \quad k = k_{,}$$

$$> 0 \quad \text{if} \quad k < k_{,}$$

$$< 0 \quad \text{if} \quad k > k_{,}$$

and

$\dot{k} = 0$ if $c = \gamma k^B - (\delta+r)k$

> 0 if $c < \gamma k^B - (\delta+r)k$

< 0 if $c > \gamma k^B - (\delta+r)k$.

These relationships are indicated in Figure 3.1. The two curves $\dot{c}$ and $\dot{k}$ divide the figure into four regions, and the behavior of c and k is indicated in each region by a pair of arrows. The two curves intersect at $(k_{,}, c_{,})$ which is the balanced growth path.

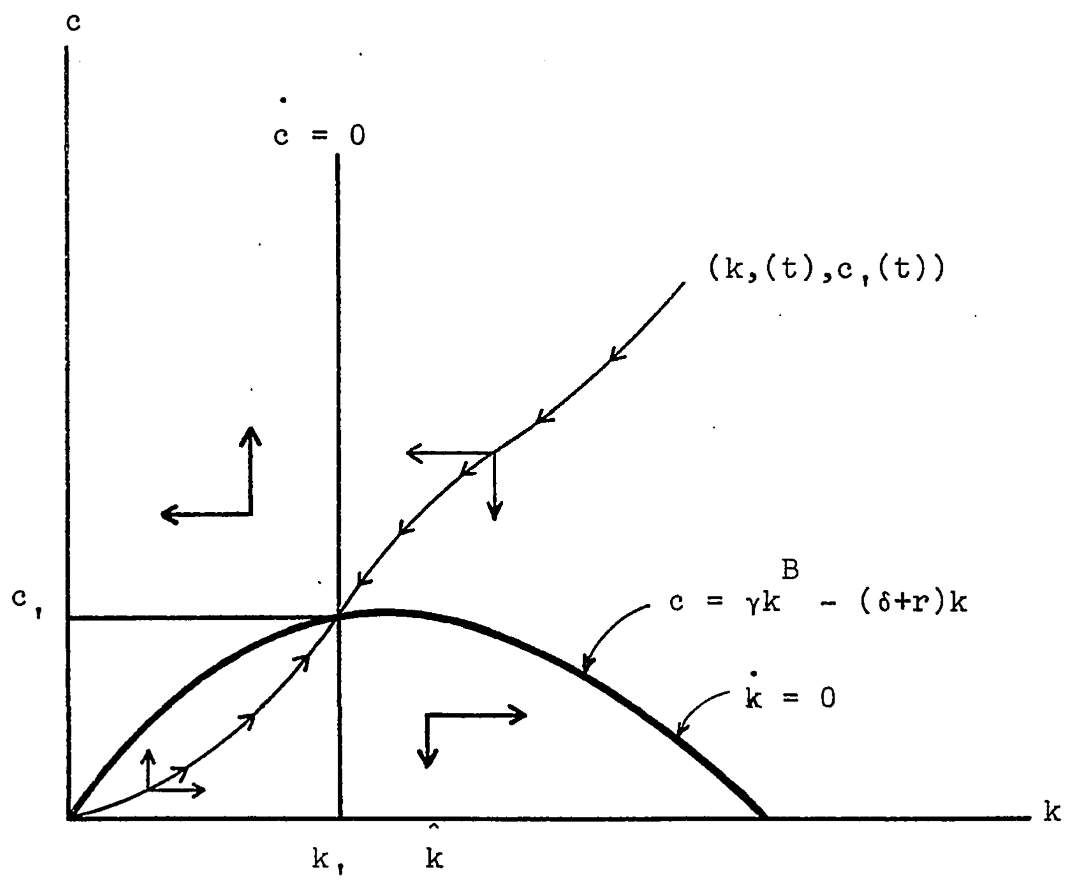

Figure 3.1. Phase diagram for problem T-3 where z = 0

The local stability of the solutions to the autonomous (since z is assumed to be zero) differential Equations 3.39 can be analyzed from the characteristic roots of the matrix of coefficients obtained by a linear expansion of these equations about the equilibrium point $(k_\prime, c_\prime)$.

Expanding about the equilibrium point $(k_\prime, c_\prime)$ one obtains:

$$\dot{c} \simeq -\frac{1}{n}[(\delta+r+\rho) - \gamma B k_\prime^{B-1}](c-c_\prime)$$

$$+ \frac{c_\prime}{n}(\gamma)(B)(B-1)k_\prime^{B-2}(k-k_\prime) \quad . \qquad (3.43)$$

By Equation 3.40 the first term vanishes and,

$$\dot{c} \simeq \frac{(.338)}{(.9)}(.285)(.6)(-.4)(3.385)^{-1.4}(k-3.385)$$

$$\dot{c} \approx -.00465(k-3.385)$$

$$\dot{k} \simeq -(c-c_\prime) + [\gamma B k^{B-1} - (\delta+r)](k-k_\prime)$$

$$\dot{k} \simeq -(c-c_\prime) + \rho(k-k_\prime)$$

$$\dot{k} \simeq -(c-.338) + (.03)(k-3.385) \quad .$$

The behavior of the system around $(k_\prime, c_\prime)$ is determined by the characteristic values of the matrix A, where A is given by:

$$A = \begin{pmatrix} 0 & -.00465 \\ -1.00 & .03 \end{pmatrix} \quad .$$

The characteristic values are determined as $x_1 = .0848$

and $x_2 = -.0548$ and two characteristics vectors are,

$$h_1 = \begin{pmatrix} -.0548 \\ 1.0 \end{pmatrix} \qquad h_2 = \begin{pmatrix} +.0848 \\ 1.0 \end{pmatrix} .$$

Since these characteristics roots are real and opposite in sign, the equilibrium point of balanced growth at (k',c') is a saddle point, the stable branch of which is labeled $(k'(t),c'(t))$ in Figure 3.1. This stable branch consists of all points that eventually reach the balanced growth equilibrium.

The path of optimal economic growth must lie along the stable branch, where given any initial level of capital per worker k_o, the unique optimal initial consumption per worker is the point on the stable branch associated with k_o. The optimal growth path is a unique segment of the stable branch, as any other path would eventually fail to satisfy the necessary conditions for an optimum involving either inflexible points in the upper left of Figure 3.1 or inferior points in the lower right of the quadrant. The stable branch is monotonic increasing, so if $k_o < k'$, then both $c'(t)$ and $k'(t)$ increase over time, moving up the stable branch to the balanced-growth equilibrium while if $k_o > k'$ the reverse is true. With a finite horizon there is an additional condition,

$$e^{-\rho t_f} q(t_f)(k(t_f) - k_f) = 0 \quad . \tag{3.44}$$

It has been shown (61) that the optimal path satisfies the "turnpike property". As the time horizon $[0,t_f]$ becomes sufficiently long, the optimal time paths for capital per worker and for consumption per worker spend an arbitrarily large portion of the time close to the balanced growth equilibrium. For example starting from its initial level k_o capital per worker moves toward $k_\prime$ and stays near there, eventually moving away from $k_\prime$ to satisfy the terminal requirement $k(t_f) = k_f$.

With the parameter values given in Table 3.7 and $B = .6$, $a = .8419$, and the growth rate of labor $r = .025$ per year, it appears from my computations that the ten year horizon does not allow the turnpike property to manifest itself for problem T-3. Consumption per worker at $t = 0$ is $c(0) = .146$ and at $t = 10.0$ has increased to $c(10) = .27$, where the equilibrium point $c_\prime = .338$ has not been reached. Likewise with $k(t)$, $k(0) = 1.0$ and $k(10) = 1.2$ where $k_\prime$, the equilibrium point is $k_\prime = 3.385$. It would appear that a time horizon of approximately 20 years would be needed to exhibit the turnpike property of problem T-3. The balanced growth solution is given by,

$$C(t) = (.338)e^{.025t}(15.0) \quad ,$$

$$K(t) = (3.385)(15.0)e^{.025t} \quad .$$

For the ten year horizon the values $k(t)$ and $c(t)$ are increasing along the stable branch of Figure 3.1. They start at $c(0) = .146$ and $k(0) = 1.0$ and within the ten year horizon do not attain the equilibrium value, $k_{,} =$ 3.385 and $c_{,} = .338$.

Problem T-3 will be considered again in a suboptimization procedure with a two-sector model that is treated in Section C.

Jorgenson (38), Sengupta (64) and Goodwin (30) considered the idea of a dual economy framework. The economic system may be divided into two sectors, the advanced (manufacturing) and the backward (agricultural) such that the production in the former is a function of labor and capital with constant returns to scale, whereas in the latter sector, production is a function of land and labor with diminishing returns to scale.

The Jorgenson model of a dual economy in its developing phase may be summarized as follows:

a. The development of the advanced sector, also called manufacturing, is possible only if an agricultural surplus eventually emerges in the backward, also called agricultural sector. If no such surplus comes into existence, the entire economy remains stagnant, producing only food and other products of the backward economy.

b. When the output of the agricultural sector attains and then exceeds the minimum subsistence level of food consumption necessary for population to grow at its maximum rate, an agricultural surplus emerges. As a result, total population (i.e. labor) $P(t) = P(0)e^{\varepsilon t}$ grows at the maximum rate of net reproduction and hence, a part of the labor force may be freed from the agricultural sector to produce goods in the advanced sector.

c. It is assumed that all income arising in the agricultural sector either as wages to labor or rent to landowners are entirely consumed while the output of the advanced sector ($X(t)$) is partly consumed ($X_c(t)$) (both directly and indirectly) through trading for food produced in the agricultural sector and partly invested ($I(t)$). Capital accumulation $\dot{K}$ which is possible only in the advanced sector, is defined as investment ($I(t)$) less depreciation $\delta K(t)$ where δ is the constant rate of depreciation.

$$X(t) = X_c(t) + I(t) = X_c(t) + \dot{K}(t) + \delta K(t) \quad .$$

d. The production functions for the agricultural $Y(t)$ and manufacturing $X(t)$ sectors are assumed to be of the Cobb-Douglas form with neutral technical changes.

$$Y(t) = e^{\alpha t} A^{1-B} \quad ,$$

$$X(t) = e^{\lambda t} K^{\sigma} M^{1-\sigma} \quad ,$$

where α, B, λ and σ are known estimated from the sectors in question. Total population $P(t)$ is made up of agricultural labor A(t) and manufacturing labor M(t).

Given the agricultural production function, the required rate of growth in the agricultural labor force necessary to maintain the growth of the agriculture surplus is computed in the model as:

$$A(t) = P(0)e^{(\frac{\varepsilon-\alpha}{1-B})t} .$$

Since the total population is growing at the maximum rate, the size of the manufacturing labor force is given by:

$$M(t) = P(t) - A(t) = P(0)[e^{\varepsilon t} - e^{(\frac{\varepsilon-\alpha}{1-B})t}] .$$

One may obtain an expression for the aggregate consumption C(t) as,

$$C(t) = Y(t) + X_c(t) .$$

By using the production functions in the two sectors and substituting into the previous equation, we obtain the differential equation

$$C(t) = e^{\alpha t}[P(0)e^{(\frac{\varepsilon-\alpha}{1-B})t}]^{1-B}$$

$$+ e^{\lambda t}K^{\sigma}[P(0)(e^{\varepsilon t} - e^{(\frac{\varepsilon-\alpha}{1-B})t})]^{1-\sigma} - \dot{K}(t) - \delta K$$

$$C(t) = P(0)^{1-B}e^{\varepsilon t} + P(0)^{1-\sigma}e^{\lambda t}K^{\sigma}[e^{\varepsilon t} - e^{(\frac{\varepsilon-\alpha}{1-B})t}]^{1-\sigma}$$

$$- \dot{K}(t) - \delta K .$$

Now we may formulate an index of performance either with a utility function of the argument C(t) or a disutility function describing deviation from a known desired time path. For this case let us consider the former:

$$\text{minimize } J = \int_{o}^{t_f} L(C(t))e^{-\rho t}dt + \frac{p}{2}(K(t_f) - K_f)^2$$

where $K(0) = K_o$ and $K(t_f) = K_f$ represent boundary condition for the problem where K_f may be computed from a given target growth rate. Sengupta (64) noted that this problem was too complicated and nonlinear to solve explicitly analytically. He analyzed the problem in various cases using linear approximations to the actual problem.

The previous formulation is isomorphic to problem T-3 and the computational procedure to numerically solve it is identical to that used in solving T-3. I make no computations, but merely point out the similarity of the two problems.

C. Two-Sector Optimal Growth Models

To begin the study of the control problem applied to two-sector models, consider the model of development and planning for India which was formalized by Mahalanobis (48). It is studied as an indication of how one might proceed with other more complex models. The model distinguishes two sectors, one producing investment goods and the other producing consumption goods on the assumption of a closed economy. The increase of real national output depends on the allocation of investment to each sector. The main policy problem is how to determine the optimal allocation of investment between the two sectors under alternative planning horizon and various sets of values of the output-capital coefficients.

The two sector model may be specified in continuous form as follows:

$$\dot{I}(t) = \lambda_i B_i I(t) \quad , \tag{3.45}$$

$$\dot{C}(t) = \lambda_c B_c I(t) \quad , \tag{3.46}$$

$$\lambda_i + \lambda_c = 1 \quad , \tag{3.47}$$

$$Y(t) = C(t) + I(t) \quad , \tag{3.48}$$

$$I(0) = I_o \quad , \quad C(0) = C_o \quad . \tag{3.49}$$

The variables are defined as,

$I(t)$ = Investment goods at time t,

$C(t)$ = Consumption at time t,

$Y(t)$ = Income at time t,

λ_i = Proportion of total investment allocated to produce investment goods,

λ_c = Proportion of total investment allocated to produce consumption goods,

B_i = Output-capital ratio for investment goods,

B_c = Output-capital ratio for consumption goods.

As indicated, the policy problem which this model is to help solve is that of providing a means to compute the proportion of total investment which should be allocated to produce investment goods, λ_i in order to maximize aggregate income $Y(t_f)$, given the planning horizon $[0,t_f]$.

The output-capital ratios are assumed known and based on production situations in each sector and given by $B_i = .2$ and $B_c = .3$ (27). From Equations 3.45 to 3.49 I(t), C(t) and Y(t) can be determined in terms of λ_i and t. Once the planning horizon $[0,t_f]$ has been specified the necessary condition for a maximum of $Y(\lambda_i;t_f)$,

$$\frac{d}{d\lambda_i} Y(\lambda_i;t_f) = 0 \quad , \tag{3.50}$$

allows one to compute the value $\hat{\lambda}_i$ which gives the maximum value of $Y(\lambda_i;t_f)$. The meaning of $\hat{\lambda}_i$ is the proportion of total allocatable investment to be made in investment goods to maximize income $Y(\lambda_i;t_f)$ in t_f years. In the above formulation the implicit welfare function includes only one element, the maximization of $Y(\lambda_i;t_f)$.

The preceding two-sector model can be linked to control problem T-2 or T-3 in the following manner. Add Equations 3.45 to 3.46 and use the time derivative of Equation 3.48 and Equation 3.47 to obtain the following differential equation,

$$\dot{Y}(t) = [\lambda_i B_i + (1 - \lambda_i)B_c]I(t) \tag{3.51}$$

where $0 \leq \lambda_i \leq 1$. Let $B(\lambda_i) = \lambda_i B_i + (1 - \lambda_i)B_c$ and from Equation 3.48, one obtains the differential equation,

$$\dot{Y}(t) = B(\lambda_i)(Y(t) - C(t)) \quad . \tag{3.52}$$

Define $c(t) = C(t)/L_o e^{rt}$ as consumption per worker where L_o is the initial labor force and r is the growth rate of labor. Form the integral functional,

$$\text{minimize } J = \int_0^{t_f} (c(t) - c^*(t))^2 dt \quad , \tag{3.53}$$

where $c^*(t)$ is a known desired consumption per worker trajectory over $[0,t_f]$. Then one may specify the boundary conditions, where Y_f is a terminal target output as,

$$Y(0) = Y_o \quad , \quad Y(t_f) = Y_f \quad . \tag{3.54}$$

This extension of the Mahalanobis model has two control variables $c(t)$ and $\lambda_i(t)$ and one state variable $Y(t)$. Classify this model as problem T-4.

$$\text{minimize } J = \int_0^{t_f} (c(t) - c^*(t))^2 dt \tag{3.55}$$

subject to:

$$\dot{Y}(t) = B(\lambda_i)(Y(t) - c(t)L_o e^{rt})$$

$$Y(0) = Y_o \text{ and } Y(t_f) = Y_f \qquad 0 \leq \lambda_i \leq 1 \quad .$$

One notes that there are two parts to the economic meaning of the optimization of this two-sector problem.

1. To determine the optimal allocation between sectors, $\lambda_i(t)$, over the planning horizon.
2. To determine the optimal allocation between consumption and production subject to the desired consumption and the dynamic constraints and boundary conditions of the system.

One approach to consider this kind of problem is by a decomposition procedure. Rather than treat $\lambda_i(t)$ as a control variable in problem T-4, one may choose a value for λ_i to optimize Y(t) as in the original Mahalanobis problem or by any other subunit optimization procedure where $\lambda_i(t)$ is the decision variable. One then obtains various values of λ_i over subintervals of the planning horizon. That is to say λ_i is constant over subintervals of $[0,t_f]$ such that,

$$
\begin{aligned}
\lambda_i &= Z_1 && t \in T_1 \\
&= Z_2 && t \in T_2 \\
&\vdots && \vdots \\
&= Z_n && t \in T_n
\end{aligned}
$$

where $T_1, T_2, T_3, \ldots, T_n$ form a partition of $[0,t_f]$ and $Z_1, Z_2, \ldots Z_n$ are constants such that

$0 \leq Z_j \leq 1$ for $j = i,2,\ldots,n.$

Once the Z_j have been chosen, $B(\lambda_i)$ is completely determined. $B(\lambda_i) = Z_j B_i + (1-Z_j)B_c$ for $t \in T_j$. This interpretation gives a step function time varying output-capital ratio, $B(\lambda_i)$, and with these modifications the problem T-4 is a slight extension of problem T-2 as treated in Section B.

If the performance functional, Equation 3.53, is given in terms of utility,

$$\text{maximize } J = \int_{o}^{t_f} L(c(t),Y(t))dt \quad . \qquad (3.56)$$

and $B(\lambda_i)$ is computed as indicated previously, then the two-sector problem reduces to the T-3 problem as treated in Section B. Classify this two-sector utility problem as T-5.

The optimal solution of this modified two-sector model T-4 may not be identical to the solution of the T-4 problem where $\lambda_i(t)$ is treated as a control variable. However trade offs may be made between the subunit and overall objective functions such that a reasonable approximation is attained. Here the modified control problem solution is optimal consistent with the subunit decisions concerning $\lambda_i(t)$.

If there is a central planning agency at the national level for a country it may not be the most efficient for the agency to make all detailed decisions concerning all the controls. The central agency may have a limited knowledge about the detailed parameters at the subunit (regional or sectorial) levels, particularly when some of the parameters are subject to dynamic shifts.

One could however, visualize two alternative ways of synthesizing subunits into a single national policy model, assuming each subunit appropriately defined can be regarded as a decision-making unit. One is to specify a team decision for national policy problems so that the optimizing considerations of all the subunits are incorporated in the one national model. An example for a simple case is the two-sector model T-4 where all the subunit decisions are made within the model, the optimal time paths for $\lambda_i(t)$, Y(t) and c(t). Alternatively, one can specify a suboptimization or multiphase decision model at the national level, where the various subunits form different phases. The central decision making agency itself may form one phase in the sequential scheme of the decision making.

If each subunit is required to fulfill a part of the national goal and also a subunit goal which is specific to the unit itself, care should be taken to ensure that the controls chosen by different subunit policy makers are compatible among themselves and in relation to the national targets set up. In the general case the team decision problem becomes one of a nonlinear time-staged programming problem or a control problem with large dimensions. In view of the data requirements it appears that the formulation of a detailed multisubunit growth model is a very difficult task in most countries (27). Also the cost and the numerical difficulties involved in the computation increases rapidly as the number of subunits increase. The computational difficulty of solving control problems seems to increase rapidly with the number of control variables.

Chenery (18) and Sengupta (64) suggest as an alternative the procedure as mentioned before of a suboptimization decision problem in a multiunit framework. To emphasize the idea of sequential planning by stages, one may start in the first stage with a dynamic macroeconomic decision problem at the national level, an aggregate growth model with a long planning horizon of ten to fifteen years.

At the next stage one considers problems of optimal decision making at subunit (possibly sectorial) levels. An objective function different from that in the first stage could be selected at this stage with a short planning horizon of three to five years. Any deviation of the observed solution from the planned targets at the end of each short planning period in the second stage could be utilized to revise the initial first-stage decision and perhaps update the general model. This revised first-stage decision could then be used in the second-stage model to get an improved decision for the next time horizon of the subunit. This multistage suboptimization procedure could be extended into additional units. Problem T-4 simply involves subunit decisions on $\lambda_i(t)$, the optimal allocation as between the two sectors and $c^*(t)$, the desired consumption per worker trajectory. This procedure allows one to work with a general control problem less difficult than the one generated by the team decision approach. However the solution to the suboptimization form of the problem is not optimal in the sense of the team decision problem since it allows for compromises and trade offs between the general and subunit objective functions. Changes made in the subunit decision variables are reflected in the value of the general objective functional.

As a numerical example of how one might proceed with this suboptimization process, consider the model T-4. Suppose that $B_i = .2$ and $B_c = .3$ are determined from the sectors in question. The allocation ratio $\lambda_i(t)$ is chosen for subintervals of a 15 year horizon by an independent suboptimization process as mentioned and the values of $\lambda_i(t)$ are given as:

$$\lambda_i = Z_1 \quad \text{if} \quad 0 \leq t \leq 5,$$

$$= Z_2 \quad \text{if} \quad 5 < t \leq 10,$$

$$= Z_3 \quad \text{if} \quad 10 < t \leq 15,$$

$$0 \leq Z_j \leq 1 \qquad j = 1,2,3 \quad .$$

The output-capital ratio is defined over [0,15] as,

$$B(\lambda_i) = \lambda_i B_i + (1-\lambda_i)B_c \quad .$$

There are two subunit decisions in this example, one of choosing λ_i has been made; the other is that of choosing $c^*(t)$. The variable $c^*(t)$ represents the desired consumption per worker trajectory. Suppose that a subunit of the economic unit in question, say the businessmen, wish a given desired path $c^*(t)$. Now the central planning authority can take as given $\lambda_i(t)$, the allocation ratio between sectors, and $c^*(t)$, the desired consumption per worker trajectory, and Y_f the desired final output of the complete unit, and solve the following problem,

$$\text{minimize } J = \int_o^{t_f} (c(t) - c^*(t))^2 dt$$

$$\text{subject to} \quad \dot{Y}(t) = B(\lambda_i)(Y(t) - c(t)L_o e^{rt})$$

$$Y(0) = Y_o \quad \text{and} \quad Y(t_f) = Y_f \quad .$$

which is a slight extension of problem T-2. Changes made in $c^*(t)$ or $\lambda_i(t)$ are reflected in J. If J is not small enough then compromises from subunit optimal values must be made in either $c^*(t)$ or $\lambda_i(t)$ and then J recomputed.

Thus by a series of trade offs from the subunit optimal values the overall "best" can be obtained for the complete unit. The "best" is measured by means of the smallest functional value J consistent with the submitted current subunit control values of $c^*(t)$ and $\lambda_i(t)$. To illustrate how c(t) and J change with respect to λ_i, the optimal control c(t) and functional value were computed for three sequences of values of λ_i with $c^*(t)$ held as $c^*(t) = 9.0 + .5t$, a subsistence term plus a linear time trend. The sequence of λ_i for this illustration were not selected as optimizing subunit values but simply values close to its maximum or minimum with switches between these values and the value $\lambda_i = 1/2$ for a comparison.

1. Equal allocation $\lambda_i = 1/2$ for $0 \leq t \leq 15$.
2. Low $\lambda_i = .1$ for $0 \leq t \leq 10$ then high $\lambda_i = .9$ for $10 < t \leq 15$.
3. High $\lambda_i = .9$ for $0 \leq t \leq 10$ then low $\lambda_i = .1$ for $10 < t \leq 15$.

The parameters used for the computation are $Y_o = 100.0$, $Y_f = 212.$ (which allows for a 5% growth rate per year on Y(t)), $t_o = 0$, $t_f = 15.0$, $L_o = 10.0$, $r = .01$ and penalty constant is 3.0. This computational example is given simply as an illustration of how one might use the suboptimization procedure and how certain of the controls are related. The value of J reflects how the subunit decisions affect the complete unit model.

One notes from Table 3.15 that sequence 2 has clearly the smallest J value. If that J value were not small enough, then compromises would have to be made in the subunit optimal values, c*(t), $\lambda_i(t)$ or possibly the terminal target constraint $Y(t_f) = Y_f$ and the general model recomputed. Also J can be computed as a function of the switching time t_1. Various computations can be made to determine the switch time which gives a minimum J.

Table 3.15. Consumption per worker c(t) for three sequences of values of the allocation ratio λ_i

1.	2.	3.
λ_i=1/2 $t\varepsilon[0,5]$	λ_i=.1 $t\varepsilon[0,5]$	λ_i=.9 $t\varepsilon[0,5]$
=1/2 $t\varepsilon[5,10]$	=.1 $t\varepsilon[5,10]$	=.9 $t\varepsilon[5,10]$
=1/2 $t\varepsilon[10,15]$	=.9 $t\varepsilon[10,15]$	=.1 $t\varepsilon[10,15]$

t	c(t)	c(t)	c(t)
0.00	6.09	7.31	4.04
1.80	8.01	8.50	7.39
3.60	9.57	10.0	9.05
5.25	11.0	11.2	10.5
7.05	12.1	12.3	11.7
9.00	13.2	13.3	12.9
10.80	14.2	14.3	13.8
12.50	15.1	15.2	14.9
13.65	15.8	15.8	15.6
15.0	16.4	16.5	16.3
	J=17.639	J=9.50	J=33.43

A more general two-sector model will now be considered which generalizes the neoclassical growth model of Section B by allowing for two sectors using different techniques of production. No computation will be performed on this and extensions of this model. Rather it will be indicated how the model could be decomposed similar to the procedure for problem T-4 and thus solved by identical computing procedures as was done with problem T-2 and T-3. In this general two-sector model we are not limited to linear production relationships. One sector produces a homogeneous capital good and the other a homogeneous consumption good.

Let $Y_c(t)$ be the output of the consumption good at time t, and $Y_i(t)$ be the output of the investment good at time t, GNP at time t, valued in terms of the consumption good is

$$Y(t) = Y_c(t) + pY_i(t) \quad , \tag{3.57}$$

where p is the price of the investment good in terms of the consumption good.

Each sector produces its output using two factors of production, capital and labor, as determined by the production functions

$$Y_j = F_j(K_j, L_j) \quad , \qquad j=c,i \tag{3.58}$$

where $K_j(t)$ is the capital employed in sector j, and $L_j(t)$ is the labor employed in sector j. Assume that each of the production functions $F_j(.,.)$ satisfies the neoclassical assumptions represented by Equation 3.6 and Equation 3.7. Also, the production functions exhibit no externalities in that the output of one sector does not depend directly upon the output or input of the other sector.

The factors of production are homogeneous and can be freely shifted between sectors. Assuming both factors are fully employed, then one has

$$K_c(t) + K_i(t) = K(t) \quad , \tag{3.59}$$

$$L_c(t) + L_i(t) = L(t) \quad ,$$

where K(t) is the aggregate stock of capital, and L(t) is the total labor force available at time t. The total capital stock is augmented by investment and subject to depreciation at the constant rate δ,

$$\dot{K} = Y_i - \delta K \tag{3.60}$$

while the labor force grows exponentially,

$$\dot{L} = rL \quad . \tag{3.61}$$

The model can be reformulated in terms of per worker quantities.

$$\frac{Y_c}{L_c} = f_c(k_c) = F_c(\frac{K_c}{L_c}, 1) \quad , \tag{3.62}$$

$$\frac{Y_i}{L_i} = f_i(k_i) = F_i(\frac{K_i}{L_i}, 1) \quad .$$

The variables k_c and k_i are sectoral levels of capital per worker and l_j is the proportion of the labor force allocated to sector j.

$$l_j = \frac{L_j}{L} \geq 0 \quad \text{for} \quad j=c,i$$

$$k_j \geq 0 \quad \text{for} \quad j=c,i$$

and

$$l_c + l_i = 1.0 \tag{3.63}$$

Consumption per worker is given by the equation,

$$y_c = \frac{Y_c}{L} = l_c f_c(k_c) \quad . \tag{3.64}$$

Investment per worker is,

$$y_i = \frac{Y_i}{L} = l_i f_i(k_i) \quad . \tag{3.65}$$

Gross National Product per worker in terms of consumption goods is thus

$$y = y_c + p y_i \quad , \tag{3.66}$$

and aggregate capital per worker in the economic unit is:

$$k = \frac{K}{L} = k_c l_c + k_i l_i \quad , \tag{3.67}$$

so from Equation 3.60 one obtains the differential equation,

$$\dot{k} = y_i - (\delta + r)k \quad . \tag{3.68}$$

The problem of optimal control for the two-sector model is then the problem of choosing time paths $\{l_i(t),\ l_c(t),\ k_i(t),\ k_c(t)\}$ such that J is a minimum, where

$$J = \int_o^{t_f} -e^{-\rho t} L(y_c)dt + \frac{p}{2}[k(t_f) - k_f]^2 \tag{3.69}$$

subject to the constraints:

$$\dot{k} = y_i - (\delta + r)k \tag{3.70}$$

$$k(t_o) = k_o \quad \text{and} \quad k(t_f) = k_f$$

$$y_c = l_c f_c(k_c)$$

$$y_i = l_i f_i(k_i)$$

$$l_i + l_c = 1.0$$

$$k = k_i l_i + k_c l_c$$

$$k_c,\ k_i,\ l_i,\ l_c \geq 0$$

and piecewise continuous where k is the state variable; l_i, l_c, k_i and k_c are the control variables and $f_c(.)$, $f_i(.)$ and $L(.)$ are given strictly concave functions; t_o, ρ, δ, r, k_o, k_f and t_f are given parameters.

If one determines l_i by a subunit optimization procedure as previously mentioned over subintervals of the planning horizon $[0,t_f]$, then this two-sector problem reduces to a slight modification of the problem T-3 considered in Section B. When l_i is known, by Equation 3.63, l_c is known. Equation 3.67 can then be solved for k_i in terms of k and k_c,

$$k_i = \frac{k - k_c l_c}{l_i} \quad .$$

The preceding control problem then is identical to that considered in Section B.

$$\text{minimize } J = \int_0^{t_f} -e^{-\rho t} L((1-l_i) f_c(k_c))dt + \frac{p}{2}[k(t_f) - k_f]^2$$

subject to,

$$\dot{k} = l_i f_i \left(\frac{k - k_c(1-l_i)}{l_i} \right) - (\delta + r)k$$

$$k(0) = k_o \qquad k(t_f) = k_f$$

the state variable is $k(t)$ and the control variable is $k_c(t)$. In the absence of a subunit optimizing procedure to compute l_i, deterministic simulation can be performed to approximate the minimum of J as a function of l_i, and the switching time t_1. Approximations to the optimal solution can thus be made by repeated computation of the control problem to minimize J with respect to $l_i(t)$.

Most investments yield their benefits in the form of identifiable goods that may be marketed or withheld. The future benefits from such an investment can be measured by the output evaluated at the price at which it can all be sold, less all current production costs. But a wide class of investments yield benefits which by their very act of production, inure to a wide class of people. These individuals cannot reasonably be excluded from the benefits and, thus a price cannot be charged that will effectively discriminate between those who want service and those who do not. Water purification provides a simple example. Services derived from government investment may not be charged for, or, if they are, the rate need not correspond to their marginal usefulness to society.

The whole purpose of investment policy is to determine optimal decisions of present and future investment, and the optimal choices at different times are interrelated. One should also be concerned that future government sector investment decisions are similarly optimal.

An extension of the model just treated is a model formulated by Uzawa (72). He considered the problem of optimum fiscal policy in terms of the techniques of optimum economic growth. The model is an aggregate two-sector growth model consisting of a private and a public sector in which both labor and private capital are used to produce goods and services. Private goods may be either consumed or accumulated as capital, while public goods are all consumed.

Many countries have come to regard fiscal policy both as an instrument to achieve short-run goals and to implement long-run objectives, such as economic growth. The Ramsey theory (60) and related work (67) are based upon an economic structure similar to that of a centrally planned economy in which a central planning bureau is free to allocate the means of production, labor and capital, in whatever manner it desires. In most countries, the allocation of the means of production is not directly governed by the state authorities. Uzawa (72) supposed that the public sector could determine not only the fiscal policy but also the allocation of capital and labor between sectors and the division of private goods between consumption and investment.

The private sector comprises business firms and households. The output produced in the private sector is assumed to be composed of homogeneous quantities so that any proportion may be either instantaneously consumed or accumulated as part of the capital stock. The public sector provides the private sector with different goods and services than those it produces. Public sector goods and services are assumed to be measurable and distributed to the private sector free of cost. Capital accumulations take place only in the private sector and public goods are not accumulated. Both production processes employ capital and labor and are subject to all the neoclassical conditions as in Equation 3.7 and 3.6. The notation is similar to the previous two-sector model, where the subscript i indicates public sector and c indicates private sector. Production processes are defined as in Equation 3.58. The quantities of capital and labor in each sector are as in Equation 3.59. The output of private goods, $Y_c(t)$, is divided between consumption, $C(t)$ and investment, $Z(t)$:

$$C(t) + Z(t) = Y_c(t) \quad . \tag{3.71}$$

The accumulation of the capital is described by

$$\dot{K}(t) = Z(t) - \delta K(t) \quad , \tag{3.72}$$

where δ is the rate of depreciation and r is assumed to be exogenously given:

$$\dot{L}(t) = rL(t) \quad . \tag{3.73}$$

The utility function of the representative member of society depends upon the amount of private goods to be consumed and upon the average quantity of public goods available at each moment. Public goods are assumed to be distributed equally among the members of the economic unit. Let $L(c(t),y_i(t))$ be the utility function where $c(t)$ and $y_i(t)$ stand respectively for the quantities of per worker consumption of private and public goods. The objective functional is represented as the discounted sum of instantaneous utilities through time:

$$J = \int_0^{t_f} L(c, y_i) e^{-\rho t} dt \tag{3.74}$$

where ρ is the rate by which future utilities are compared with the present utilities.

Suppose that the public sector can determine not only the fiscal policy but also the allocation of capital and labor between sectors and the division of private goods between consumption and investment. The public sector then seeks for the feasible time paths of factor and output allocation at which 3.74 is maximized. The problem is more precisely defined as follows: Find a time path of $\{K_c(t), K_i(t), L_c(t), L_i(t), C(t), Z(t), Y_i(t)\}$ for which the functional

$$J = -\int_0^{t_f} u\left(\frac{C(t)}{L(t)}, \frac{y_i(t)}{L(t)}\right) dt \tag{3.75}$$

is minimized subject to the constraints:

$$C(t) + Z(t) \leq F_c(K_c(t), L_c(t)) \quad , \tag{3.76}$$

$$Y_i(t) \leq F_i(K_i(t), L_i(t)) \quad ,$$

$$K_c(t) + K_i(t) \leq K(t) \quad ,$$

$$L_c(t) + L_i(t) \leq L(t) \quad ,$$

$$\dot{K}(t) = Z(t) - \delta K(t) \quad ,$$

$$L(t) = L_o e^{rt}$$

with given initial $K(0) = K_o$ and terminal $K(t_f) = K_f$ where all variables are nonnegative.

Using the same notation as in the previous two-sector model with the addition of $z = Z(t)/L(t)$, per worker investment in the private sector, and omitting the time suffix and assuming full employment of all factors of production, the problem is reduced to the following:

$$\text{minimize } J = -\int_o^{t_f} L(c,y_i)e^{-\rho t}dt + \frac{p}{2}[k(t_f) - k_f]^2 \tag{3.77}$$

subject to the constraints:

$$c + z = f_c(k_c)l_c \quad , \tag{3.78}$$

$$y_i = f_i(k_i)l_i \quad ,$$

$$k_c l_c + k_i l_i = k \quad ,$$

$$l_i + l_c = 1 \quad ,$$

$$\dot{k} = z - (r+\delta)k \quad ,$$

$$k(0) = k_o \quad \text{and} \quad k(t_f) = k_f \quad .$$

The utility function, $L(c,y_i)$ is continuously twice differentiable and has positive marginal utilities L_c and L_{yi}: for all positive c and y_i and is strictly concave for all the values of c and y_i. Also the following properties hold for L:

$$L_{cc} < 0 \ , \quad L_{y_i y_i} < 0 \ , \quad L_{cy_i} \leq 0 \ ,$$

$$L_{cc}L_{y_i y_i} - L^2_{cy_i} > 0 \quad ,$$

$$\frac{L_{y_i y_i}}{L_{y_i}} - \frac{L_{cy_i}}{L_c} < 0 \ , \quad \frac{L_{cc}}{L_c} - \frac{L_{y_i c}}{L_{y_i}} < 0 \quad .$$

By combining the constraints, this control problem can be reduced to one with three control variables $\{k_i, l_i, c\}$ and one state variable, k.

$$\text{minimize } J = -\int_o^{t_f} L(c, f_i(k_i)l_i)e^{-\rho t}dt + \frac{p}{2}(k(t_f) - k_f)^2$$

subject to:

$$\dot{k} = (1-l_i)f_c\left(\frac{k-k_i l_i}{1-l_i}\right) - c - (r+\delta)k \quad ,$$

$$k(0) = k_o \quad \text{and} \quad k(t_f) = k_f \quad ,$$

$$0 \leq l_i \leq 1 \quad .$$

This problem could be computed directly as indicated in Chapter 2 using the Davidon algorithm together with sequential penalty functions to handle both the terminal constraint on k(t) and the inequality constraint on $l_i(t)$.

Another approach to solve this problem would be to decompose it into a simpler problem like T-3. Suppose that a linear relationship between f_i and f_c is assumed as

$$l_i f_i(k_i) = V l_c f_c(k_c)$$

where V is a constant. This together with the equation $l_i + l_c = 1.0$ allows one to obtain the relationship,

$$k_i = h(l_c, k_c) \quad .$$

From the equation $k_i(1-l_c) + k_c l_c = k$ a function relating k_c to l_c and k can be determined,

$$k_c = g(l_c, k) \quad .$$

Hence the control problem becomes:

$$\text{minimize } J = \int_0^{t_f} U(c,k,l_c)e^{-\rho t}dt \quad ,$$

$$\dot{k} = l_c f_c(g(l_c,k)) - c - (r+\delta)k \quad ,$$

$$k(0) = k_o \quad \text{and} \quad k(t_f) = k_f \quad ,$$

$$0 \leq l_c \leq 1 \quad .$$

Now by choosing $l_c(t)$ at values close to its maximum and minimum or by an independent suboptimization procedure where $l_c(t)$ can be determined, the control problem reduces to the problem T-3 of Section B. Various parameter changes can be made as was done in Section B to determine how J changes with respect to the constant V and the values for the $l_c(t)$ in the absence of any suboptimization procedure to choose $l_c(t)$.

Arrow and Kurz (3) consider a similar two-sector model as that of Uzawa (72) just treated. They differ in the conception of the role of public capital in the economic system. Uzawa (72) assumes that the output in each of the private and public sectors is determined by the amount of capital and labor invested in it, while Arrow and Kurz (3) assume that private output depends upon the amounts of both kinds of capital as well as of labor (one production function). A version will be briefly presented as well as how it can be reduced for computation purposes to a problem similar to T-3.

The following notation will be used:

$K_p(t)$ = total capital employed in the private sector at t,

$K_g(t)$ = total capital employed in the public sector at t,

$\tilde{k}_g(t)$ = capital per capita employed in the public sector at t,

$K(t) = K_p(t) + K_g(t)$,

$\tilde{c}(t)$ = per capita consumption at t,

$L(t)$ = labor supply at t proportional to population $P(t)$ at t,

$P(t)$ = total population at t,

$Y(t)$ = output at time t.

The output is determined by

$$Y(t) = A(t)F(K_p(t),K_g(t),L(t)) \quad ,$$

where F is a concave production function and A(t) allows for neutral technological changes.

The natural constraint is

$$\dot{K}_p(t) + P(t)\tilde{c}(t) + \dot{K}_g(t) = A(t)F(K_p(t),K_g(t),L(t)) \; .$$

The control problem can then be formulated as follows:

$$\text{minimize } J = \int_o^{t_f} e^{-\rho t} P(t)U(\tilde{c}(t),\tilde{k}_g(t))dt \quad ,$$

subject to

$$\dot{K}(t) = A(t)F(K_p(t),K_g(t),L(t)) - P(t)\tilde{c}(t) \quad ,$$

$$K(t) = K_p(t) + K_g(t) \quad ,$$

$$K_g(t) = P(t)\tilde{k}_g(t) \quad ,$$

$$K(0) = K_o \quad \text{and} \quad K(t_f) = K_f \quad ,$$

$$L(t) = L_o e^{\pi t}$$

where $U(\tilde{c},\tilde{k}_g)$ is a concave function.

Suppose we now let:

$$K_p(t) = h(t)K(t) \ , \ \text{such that} \ \ 0 \leq h(t) \leq 1$$

where h(t) is a step function defined over $[0,t_f]$. The problem can then be formulated as minimize J, where

$$J = -\int_0^{t_f} e^{-\rho t} P(t)U[\tilde{c}(t),(\frac{K(t) - h(t)K(t)}{P(t)})dt$$

$$+ \frac{PC}{2}(K(t_f) - K_f)^2$$

subject to

$$\dot{K}(t) = A(t)F(h(t)K(t),\ 1\text{-}h(t)K(t),\ L(t)) - P(t)\tilde{c}(t)$$

$$K(0) = K_o \quad \text{and} \quad K(t_f) = K_f$$

$$L(t) = L_o e^{\pi t} \quad .$$

This problem is computed identically to problem T-3 in Section B and can be solved numerically for various values of h(t) to obtain a relationship between J and h(t).

IV. GENERALIZATION OF THE COMPUTING FRAMEWORK

A. Generalization of Two-Sector Growth Models

In order to generalize the model represented in Chapter 3 as problem T-4, consider the division of the consumption good sector into three subsectors. The model represented by Equations 3.45 and 3.46 can be written as:

$$\dot{I}(t) = \lambda_i B_i I(t) \quad , \tag{4.1}$$

$$\dot{C}_1(t) = \lambda_1 B_1 I(t) \quad ,$$

$$\dot{C}_2(t) = \lambda_2 B_2 I(t) \quad ,$$

$$\dot{C}_3(t) = \lambda_3 B_3 I(t) \quad ,$$

where

C_1 = Consumption goods produced by modern factories,

C_2 = Consumption goods produced by small, family type factories,

C_3 = Services.

One may include with Equations 4.1 the following,

$$Y(t) = I(t) + C_1(t) + C_2(t) + C_3(t) \quad , \tag{4.2}$$

and

$$\lambda_i + \lambda_1 + \lambda_2 + \lambda_3 = 1 \tag{4.3}$$

where Y(t) is the aggregate output and I(t) represents investment goods. The proportion of total investment allocated to produce investment goods is λ_i, while λ_1, λ_2, λ_3 represent the proportion of total investment allocated to the three subsectors respectively. The values B_i, B_1, B_2, and B_3 are known numbers derived from the subsector or sector in question.

There are various ways in which this model can now be linked to a control problem. One way is to simply add the Equations 4.1 and use the time derivative of 4.2 to obtain the following differential equation.

$$\dot{Y}(t) = (\lambda_i B_i + \lambda_1 B_1 + \lambda_2 B_2 + \lambda_3 B_3)[Y(t) - C_1(t) - C_2(t) - C_3(t)] \tag{4.4}$$

where $Y(0) = Y_o$ and $Y(t_f) = Y_f$ are known values. One may form as a performance functional,

$$\text{minimize } J = \int_o^{t_f} [w_1(C_1(t) - C_1^*(t))^2 + w_2(C_2(t) - C_2^*(t))^2 + w_3(C_3(t) - C_3^*(t))]dt \quad , \tag{4.5}$$

where $C_1^*(t)$, $C_2^*(t)$, $C_3^*(t)$ are desired levels of consumption available in each subsector and

$$C^*(t) = C_1^*(t) + C_2^*(t) + C_3^*(t)$$

is the desired total consumption available. The w_j are known weights assigned to the deviations from the desired paths. The Equations 4.3, 4.4, 4.5 and the boundary conditions form the control problem formulation. The problem has one state variable $Y(t)$ and the seven control variables, $\lambda_i(t)$, $\lambda_1(t)$, $\lambda_2(t)$, $\lambda_3(t)$, $C_1(t)$, $C_2(t)$, $C_3(t)$. One of the λ_j can be eliminated by Equation 4.3.

A more operational formulation would be in the following modification. Consider the two-sector model in the format of problem T-4 as discussed in Section C of Chapter 3. In that model, $C = C_1 + C_2 + C_3$ and Equation 4.1 collapses into the two-sector version. From the two-sector problem select the optimal λ^*_i on the basis of minimizing the performance functional, Equation 3.53, within the given planning horizon. Also from the solution one obtains $I(t)$, $Y(t)$ and $C(t)$ at discrete time points over $[0,t_f]$. The allocation ratios between subsectors could then be chosen to secure balance with marginal proportions of consumption demand. For example, if α_1, α_2, α_3 denote the marginal propensities to consume of the three types of consumption goods and $\lambda_1 + \lambda_2 + \lambda_3 = 1 - \lambda^*_i$ is the condition of full utilization of investment, the balancing values of λ_j can be specified as:

$$\lambda_j = \alpha_j(1-\lambda^*_i)/(\sum_{i=1}^{3} \alpha_i) \quad , \quad j=1,2,3 \quad . \qquad (4.6)$$

With the λ_j values thus selected, and the $I(t)$ function known at a discrete set of time points on the basis of optimizing within the related two-sector problem, the values of $C_j(t)$ can be computed from Equations 4.1. One may either specify $C_j(t)$ at the initial time, t_o or the final time, t_f. If the former is used and the Equations 4.1 involving $C_j(t)$ may be discretized and using the known values of I(t) computed forward in time. While if the latter is used then the procedure is to discretize and move backward from t_f to t_o in time. This same procedure could be utilized for more than three subsectors, always using the optimal results from the related two-sector problem.

Consider now an intersectorial generalization of the modified Goodwin model represented in Section B of Chapter 3 as control problem T-2. One may assume an n-sector interdependent model of the dynamic Leontief type input-output scheme. Also let us assume time dependent sectorial coefficients as in problem T-4.

Denote the n component column-vector of real consumption, real national income and net investment by C, Y and I, respectively, and define the intersectoral capital-output time-dependent coefficient by B(t), where $B^{-1}(t)$ exists for all $t \varepsilon [0,t_f]$.

$$I = B(t)\dot{Y} \quad \text{where} \quad B(t) = (b_{ij}(t)) \tag{4.7}$$

$$i = 1,\ldots,n$$

$$j = 1,\ldots,n \quad .$$

The following equations define the generalized model in terms of the consumption and income vectors.

$$C = Y - I = Y - B\dot{Y} \tag{4.8}$$

$$U = \sum_{i=1}^{n} u_i = \sum_{i=1}^{n} w_i(c_i - c^*_i)^2 \tag{4.9}$$

where the performance functional is given as

$$\text{minimize } J = \int_o^{t_f} \sum_{i=1}^{n} w_i(c_i - c^*_i)^2 dt \tag{4.10}$$

and the w_i are given weights. The desired consumption time path in the ith sector is given by $c^*_i(t)$. Let the boundary conditions be given as:

$$Y(0) = Y_o \quad , \quad \text{and} \quad Y(t_f) = Y_f \quad . \tag{4.11}$$

Equations 4.8, 4.10, and 4.11 form the control problem.

One first notes that if the time-dependent intersectorial investment matrix B(t) is strictly diagonal, then the above intersectorial model decomposes into n independent control problems each of which is identical to problem T-2 except for a time varying capital-output function. The computation in this case is simply repeated solution of a problem like T-2 with terminal constraints on the final sectorial output. The problem for the ith sector is:

$$\text{minimize } J_i = \int_o^{t_f} w_i(c_i - c^*_i)^2 dt + \frac{p}{2}(y_i(t_f) - y_{f_i})^2 \quad . \tag{4.12}$$

subject to,

$$\dot{y}_i = \frac{1}{b_{ii}(t)} (y_i - c_i) \quad , \tag{4.13}$$

$$y_i(0) = y_{o_i} \quad , \quad y_i(t_f) = y_{f_i} \quad . \tag{4.14}$$

The economic meaning of this is that if the different sectors are relatively independent in the sense that the marginal capital requirements for increasing output in any sector are obtained either entirely through that sector itself or from outside the n sector system by a central planning authority, then for each sector an optimal set of time paths for y_i and c_i can be determined.

This assumption of independence may be viewed as a specific type of disaggregating the economy. The question of whether this is empirically realistic or statistically estimable is a separate issue. However from the computational point of view this assumption of independence allows for a simple means of computing the optimal trajectories for each sector.

In the case where B(t) is not strictly diagonal the problem becomes one of n state variables $y_1,\ldots,y_n$ and n control variables $c_1,\ldots,c_n$ with terminal constraints on the state variables.

Computationally this problem is a generalization of problem T-2. It requires n penalty constants and n control variables with no inequality constraints. This type of problem can be solved by the Davidon and conjugate gradient methods but computational difficulties increase with the number of penalty constants used and the number of control variables. The procedure for solution is as discussed in Chapter 2.

However from my computational experience the iterative techniques of Chapter 2 have limitations for a problem with a large number of control variables and where many penalty constants are used.

An approach to numerically solve the optimal control problem with linear inequality constraints and a large number of control variables that appears to be more operational than applying the Davidon or conjugate gradient methods with penalty functions is extensions of the discrete model developed by Bruno (9). One may generalize from the two-sector model treated in Section E of Chapter 2 to any number of activities for consumption goods and investment goods and still only one homogeneous capital good. Also treatment of the case of any number of depreciable capital goods will be considered. In both cases the technology matrix A(t) may be a known time varying matrix function.

Consider the model of one activity to produce consumption goods and two activities to produce a depreciable capital good. The notation used will be identical to that used in Section E of Chapter 2 with the addition of z_1, z_2, which represent investment per worker in activity one and two respectively. In per worker terms the problem can be formulated as follows:

$$\text{maximize } J = \int_0^{t_f} c(t)e^{-\rho t}dt \tag{4.15}$$

subject to,

$$a_{11}(t)c(t) + a_{12}(t)z_1(t) + a_{13}(t)z_2(t) + \varepsilon_o(t) = 1$$

$$(4.16)$$

$$a_{21}(t)c(t) + a_{22}(t)z_1(t) + a_{23}(t)z_2(t) + \varepsilon(t) = k(t)$$

$$\dot{k}(t) = -(r+\delta)k(t) + z_1(t) + z_2(t) \tag{4.17}$$

$$k(0) = k_o \quad \text{and} \quad k(t_f) = k_f$$

where A(t) is a known matrix function, c(t) is per worker consumption and k(t) is per worker stock of capital.

This problem, by using the maximum principle as was done in Section E, Chapter 2, reduces to finding the solution of the linear programming problem,

$$\text{maximize } H_1 = c(t) + \pi(t)(z_1(t) + z_2(t)) \tag{4.18}$$

subject to the constraints 4.16 at each discrete time point.

The dual is

$$\text{minimize } D_1 = w(t) + s(t)k(t) \tag{4.19}$$

subject to,

$$a_{11}(t)w(t) + a_{21}(t)s(t) - p_o(t) = 1 \tag{4.20}$$

$$a_{12}(t)w(t) + a_{22}(t)s(t) - p_1(t) = \pi(t)$$

$$a_{13}(t)w(t) + a_{23}(t)s(t) - p_2(t) = \pi(t)$$

where w(t), s(t) represent the real wage and rental price of capital all measured in consumption units. From linear programming theory one has,

$$z_1p_1 = z_2p_2 = cp_o = w\varepsilon_o = s\varepsilon = 0 \quad . \tag{4.21}$$

The dynamic equations that link together the various discrete time points are,

$$\dot{k} = -(r+\delta)k + z_1 + z_2 \quad , \tag{4.22}$$

$$k(0) = k_o \quad \text{and} \quad k(t_f) = k_f \quad .$$

and

$$\dot{\pi} = (r+\delta+\rho)\pi - s \quad . \tag{4.23}$$

As in Section E of Chapter 2 ρ is the time rate of discount, r is the exogenously given growth rate of labor and δ is the depreciation rate of capital.

The neighboring extremal method as considered in Section E of Chapter 2 can be implemented to solve this problem and extensions of it to ℓ activities for consumption goods and m activities for investment goods. One estimates $\pi(0)$

and with the given value $k(0)$ this allows one to solve the linear program given by Equations 4.18 and 4.16 at time point $t = 0$ and its dual Equations 4.19 and 4.20. This results in computed values for $c(0)$, $z_1(0)$, $z_2(0)$, $s(0)$, $w(0)$, $p_1(0)$, $p_2(0)$, $p_o(0)$, $\varepsilon_o(0)$ and $\varepsilon(0)$. Then using the differential Equations 4.22 and 4.23, $k(t_1)$ and $\pi(t_1)$ are computed where t_1 is the first discrete time point.

The process of computing the linear program and its dual are thus continued at each discrete time point using the differential equations to link together the time point values of $k(t)$ and $\pi(t)$ until $k(t_f)$ is determined. If $k(t_f)$ does not approximate k_f then a new value for $\pi(0)$ is considered and the procedure is repeated until $k(t_f)$ approximates k_f. The value $k(t_f)$ is dependent upon $\pi(0)$ and an interpolating procedure can be used to improve the choice of $\pi(0)$ after each iteration.

The preceding model could easily have been generalized to include ℓ activities producing consumption goods and m activities to produce homogeneous investment goods. The computational procedure would be identical to that already considered.

Consider now the general n-sector model with heterogeneous capital goods. An economy produces n+1 goods, a consumption good c, and n depreciable goods I_i with exponential depreciation rates $\delta_i (i=1,2,...,n)$. Assume labor, as before, to be growing at an exogenously fixed rate r. The notation will be the same except that subscripts will be added to the variables involving the capital goods (z, k, λ, ε, s, π, p).

The problem now becomes:

$$\text{maximize } J = \int_{o}^{t_f} c(t)e^{-\rho t}dt \tag{4.24}$$

subject to the n+1 constraints,

$$a_r(t)c(t) + \sum_{i=1}^{n} a_{ri}(t)z_i(t) + \varepsilon_r(t) = k_r(t)$$

$$r=0,1,\ldots,n \tag{4.25}$$

where $k_o=1$ and all variables c, z_i and k_r are understood to be nonnegative. There exists a differential equation for each capital good.

$$\dot{k}_i = -(r+\delta_i)k_i + z_i \quad , \tag{4.26}$$

$$k_i(0) = k_{i_o} \quad \text{and} \quad k_i(t_f) = k_{i_f} \quad , \quad (i=1,\ldots,n) \quad . \tag{4.27}$$

The Hamiltonian is formed as,

$$H = e^{-\rho t}[c(t) + \sum_{i=1}^{n} \pi_i(t)(z_i(t) - (r+\delta_i)k_i(t))] \quad , \tag{4.28}$$

where the adjoint variable is $\pi(t)e^{-\rho t}$.

One may rewrite the function to be maximized as L, where L is defined as follows,

$$L = He^{\rho t} + \sum_{i=1}^{n} \pi_i(t)k_i(t)(r+\delta_i)$$

$$= c(t) + \sum_{i=1}^{n} \pi_i(t)z_i(t) \quad . \qquad (4.29)$$

The summation term on the left is a constant at each discrete time point and is known initially and generated subsequently by the system differential Equations 4.26 and 4.27 and the following adjoint differential equations

$$\dot{\pi}_i(t) = (r+\delta_i+\rho)\pi_i(t) - s_i(t) \quad , \quad i=1,\ldots,n \quad . \qquad (4.30)$$

Hence maximizing L also maximizes H.

The primal linear program necessary to solve the control problem is then the objective function 4.29 and the constraints are Equations 4.25. The dual linear program system becomes,

$$D = w(t) + \sum_{i=1}^{n} s_i(t)k_i(t) \qquad (4.31)$$

subject to the constraints,

$$w(t)a_{o_i} + \sum_{r=1}^{n} s_r(t)a_{r_i} - p_i(t) = \pi_i(t) \quad ,$$

$$i=1,\ldots,n \quad . \qquad (4.32)$$

$$a_o w(t) + \sum_{r=1}^{n} a_r(t)s_r(t) - p_o = 1 \quad . \qquad (4.33)$$

The variables $w(t)$, $s_r(t)$, $(r=1,\ldots,n)$ are the wage rate and the rental price of the types of capital measured in consumption units. This model has n sectors and a single technique for producing the consumption good.

The values $k_i(0)$ are given and one then approximates $\pi_i(0)$ $(i=1,\ldots,n)$. With this information the primal linear program whose objective function is given by Equation 4.29 and constraints by Equations 4.25 and the dual 4.31, 4.32, and 4.33 are solved. Using the differential Equations 4.26 and 4.30 together with the boundary conditions 4.27, $\pi_i(t)$ and $k_i(t)$ are stepped up in time and the process is repeated until $k_i(t_f)$ $(i=1,\ldots,n)$ are computed. These values are compared to k_{i_f} and if all the values are not within a given tolerance of $k_i(t_f)$, then the $\pi_i(0)$ $(i=1,\ldots,n)$ are rechosen and the computation repeated until the boundary conditions are approximately satisfied.

One notes that this numerical procedure to solve the previously mentioned control problem does not require a constant technology matrix, but allows for a time varying matrix function. The computational procedure requires a process of altering the initial values of $\pi_i(0)$ until the boundary conditions $k_i(t_f) = k_{i_f}$ are satisfied.

B. Simulation of Optimal Trajectories

Another approach by which one may numerically solve and study optimal economic growth problems relates to simulation, both stochastic and deterministic. The objective of the simulated optimization approach is to develop efficient, economical techniques for locating improved but not necessarily optimum solutions to models where other optimization techniques cannot be realistically applied or are too costly to utilize.

A great deal of literature on control theory, feedback and sensitivity analysis is relevant to this problem, for example Fox, Sengupta and Thorbecke (27), Sen (63), Hestenes (32), Theil (69), Naylor, Wertz and Wonnacott (55), Naylor et al. (54), Fromm (28), and Fromm and Taubman (29). Fromm and Taubman (29) have applied the technique of simulation via repeated solution of an economic model to compute the utility of alternative policy actions for evaluating the relative desirability of a set of monetary and fiscal policy actions. Naylor, Wertz and Wonnacott (55) used stochastic simulation to compare the stability of various policy actions by statistical techniques.

At least four general alternatives are available to economic policy makers for evaluating the effectiveness of their decisions involving economic policies. First it may be possible to perform controlled experiments with the given economic system, where the system may be a firm, an industry, or the economy of a country. Usually institutional, political and other practical constraints make this alternative impossible in the case of an industry or the economy as a whole, and difficult in the case of a firm. Even where actual experiments may be carried out it is unlikely that the relevant variables can be held constant to obtain meaningful comparisons of policy alternatives.

Second, one may use an objective functional as was done in the earlier work to determine an index of performance on the economic system. Optimizing the functional subject to the equations describing the system gives a measure of effectiveness of the optimal policy with respect to the index chosen. Parameter variations can then be made using the functional value as an indicator of sensitivity measure as was previously done.

Third, if cross-section data are available over time, it may be possible to perform a type of ex post experiment with an economic system.

Fourth, when controlled experimentation is impossible or impractical and cross-section data is unavailable, then the policy maker may use the following alternative. He may formulate and estimate the parameters of the model of the given system relating the endogenous variables of the system to the exogenous variables and policy instruments or controls.

If the model consists of a large number of simultaneous, nonlinear differential or difference equations possibly with stochastic error terms included then analytical techniques exist in theory only. When this is the case one must resort to numerical analysis techniques which were treated earlier and/or simulation to evaluate alternative economic policies.

Simulation may be defined as a numerical procedure for conducting experiments on a digital computer with certain types of mathematical models describing the behavior of an economic system over extended periods of time (54). The simulation may either be stochastic in which random variables are involved or deterministic where parameter modifications are considered. For example in the optimal economic growth problem deterministic simulation may involve experimentation with various feedback relationships or possible ways of simplifying a complex model as was done in Section C of Chapter 3. Stochastic simulation allows for stochastic error terms to be included in possible feedback relationships or in production processes. The principle difference between a simulation experiment and a "real world" experiment is that with simulation the experimentation is conducted with a model of the economic system rather than the actual economic system itself.

A question of interest is how does the optimal solution computed from a growth model with a performance functional based on utility of consumption compare with the consumption path computed from various runs of a simulated system based on some sort of feedback relation either deterministic or stochastic. Comparisons may be considered either on magnitude of consumption, C(t), or its variability, or of the computed utility of the consumption. Experimentation based on the computed optimal paths from problem T-3 will be compared with various simulated results. One wishes to find the combination of parameter values or factor levels at which the response variable is maximized to optimize some process, in this case the objective functional.

For the first experimentation, consider the discretized version of problem T-3, where the parameters are as defined in Section B, Chapter 3. The objective function becomes

$$\text{maximize } J = \sum_{i=1}^{T-1} \frac{1}{(1+\rho)^i} \frac{1}{1-n} (c_i)^{1-n} - \frac{\rho c}{2} [K(T) - K_T]^2$$

$$+ [\frac{1}{1-n} c_o^{1-n} + \frac{1}{(1+\rho)^T} \frac{c_T^{1-n}}{1-n}](.5) \qquad (4.34)$$

subject to the difference equations,

$$K_{i+1} = Y_i - C_i + (1-\delta)K_i \, , \qquad (4.35)$$

$$Y_i = (1+g)^i a K_i^B \qquad (4.36)$$

where K_o and K_T are given values and $g = r(1-B) + z$ and pc is a positive penalty constant.

An example of the use of deterministic simulation will be treated for the problem T-3 with respect to a feedback relationship. The parameter values for the model are given as follows:

$B = .6$, $n = .9$, $\delta = .05$, $K_T = 24.0$ (a 5% rate of growth on output Y_i), time horizon [0,10], rate of labor growth, $r = .025$, neutral rate of technical change, $z = .01$ and penalty constant, $pc = 3.0$.

A feedback relationship of the form

$$C_i = g(Y_i, Y_{i-1}, \ldots, Y_{i-j}) \qquad (4.37)$$

is considered, where C_i is consumption at the ith period and Y_i is the output of the economic system at the ith period. The first case treated will be linear with no intercept term of the form

$$C_i = \alpha_1 Y_i \quad (0 < \alpha_1 < 1) \qquad (4.38)$$

$$C_i = \alpha_1 Y_{i-1} \quad .$$

One notes that Equation 4.37 and 4.38 are exact relationships and have no stochastic error terms. The parameter α_1 is then allowed to assume various values and for each value the relationship 4.38 is substituted into the difference equation system 4.35 and 4.36. Thus K_i, C_i, Y_i for $i=0,1,2,\ldots,T$ can be computed.

From the computed values the objective function 4.34, the utility of alternative policy actions, is determined for each modified value of α_1. Results for the discretized T-3 problem follow in Table 4.1.

A regression of the optimal time path data computed in Section B of Chapter 3 gives the relationship between C_i and Y_i as $C_i = .674\ Y_i$ with a multiple determination coefficient, R^2, of .78. It is interesting to note that the simulated value of $\alpha_1 = .670$ yields the objective function J extremely close to the optimal computed functional value of J = 98.18. For the values of α_1 greater than .70 the computed value of

Table 4.1. Objective function values for different choices of the feedback constant α_1 (deterministic simulation)

$C_i = \alpha_1 Y_i$		$C_i = \alpha_1 Y_{i-1}$	
J	α_1	J	α_1
91.14	.62	83.17	.63
96.66	.64	91.76	.65
98.04	.67	96.65	.67
96.09	.68	98.03	.69
90.56	.70	96.05	.71
81.90	.72	90.87	.73
70.39	.74	82.67	.75

$Y(t)$ is much below the target value Y_T. For the model T-3 the neutral technological growth parameter would have to be greater than this run value of $z = .01$ to allow a coefficient value (marginal propensity to consume) greater than $\alpha_1 = .66$ to be an optimal value.

For the lagged relationship $C_i = \alpha_1 Y_{i-1}$ a least square fit of the optimal time path data computed in Section B of Chapter 3 gives the relationship $C_i = .694\ Y_{i-1}$ for the ten year horizon. The multiple determination coefficient is .80 and the residual variance equals .26. Again the feedback simulation value of the parameter $\alpha_1 = .69$ gives an extremely close simulated value to the feedback coefficient obtained by the regression of C_i on Y_i from the optimal time path data computed from the control theory algorithms.

A reason for including a stochastic disturbance term in the model is that one may replicate the simulation experiment for given stochastic parameter specifications and then construct confidence intervals and make probabilistic inferences about the differences in the effects of alternative parameter choices. Without the inclusion of these disturbance terms, one can say little about the statistical precision of the inferences made about the effectiveness of parameter choices on the basis of simulation experiments. Also wars, foreign competition, labor strikes, and national disasters are factors which might affect national income and consumption but which may not be subject to prediction and control by the policy makers.

Factorial experimental designs and multiple comparison techniques are relevant to analyzing simulation data (62). For example two parameters of interest may be considered at five levels each. If one requires a complete investigation, including main effects and interaction of all orders this requires 25 cells. Replication within each cell can be made a given number of times. Less than a complete investigation will require less cells and thus less computer time and simulated data. One then searches for the factor levels at which the objective function is maximized.

The control problem feedback simulation experiments on problem T-3 which were conducted consisted of four runs, one for each parameter specification. In each run the economy was simulated for a period equal to ten time units, and then J was computed. The simulation was replicated ten times using the given relationship of the feedback function together with a stochastic disturbance term. The feedback function used was $C_i = \alpha_1 Y_i + u_i$ where u_i are normally and independently distributed with mean 0 and variance equal .28. The results are summarized in Tables 4.2 and 4.3. The pseudorandom numbers generated were independently computed for each run and for each parameter modification.

From the data in Table 4.3 the F value is computed as

Table 4.2. The objective function value, J, for different feedback parameter values with a stochastic error term in the feedback relationship

$C_i = \alpha_1 Y_i + u_i$		u_iN.I.D. (0,.28)	
$\alpha_1 = .63$	$\alpha_1 = .67$	$\alpha_1 = .72$	$\alpha_1 = .76$
96.46	97.68	87.08	61.22
95.99	95.41	84.20	22.09
93.90	93.62	82.23	88.45
61.52	93.90	97.62	68.71
85.33	87.18	36.43	34.49
68.77	81.98	87.82	52.28
77.45	85.86	67.76	56.22
62.10	91.16	74.84	65.95
64.65	96.89	83.13	45.83
70.30	92.28	88.30	51.28
$\bar{J}_1 = 77.65$	$\bar{J}_2 = 91.59$	$\bar{J}_3 = 78.94$	$\bar{J}_4 = 54.65$

Table 4.3. Statistics for one-way analysis of variance

Source of Variation	Sum of Squares	d.f.	Mean Square
Between	7098.82	3	2366.27
Error	7727.8	36	214.66
Total	14,826.62		

$F = 11.023$. The null hypothesis, H_o, is that the population means for the objective function values for the different parameter values are all equal. By employing the F statistic, the decision rule for accepting or rejecting H_o becomes:

if $F \geq F_\alpha$ (3, 36), then reject H_o,

otherwise accept H_o, where α is the significance level.

The value of $F_{.05}$ (3, 36) is 3.28 and for $F_{.01}$ (3, 36) is 4.40, hence the data generated by the simulated experiment do not support the null hypothesis. Rejection of the null hypothesis is made at both the .05 and .01 significance level.

Having rejected the hypothesis that the objective function value associated with each of the four feedback relations is the same, one may now consider multiple comparisons between the feedback relationships.

Tukey's method (62) will yield simultaneous confidence intervals for the differences between all pairs of means. With a 95% probability, all of the following confidence intervals are true.

Let J_{ij} be the functional value of the ith replication of the jth parameter modification and $\bar{J}_j$ be the mean of the jth modification.

$$(EJ_j - EJ_e) = (\bar{J}_j - \bar{J}_e) \pm q_{k,p}\sqrt{\frac{MS_e}{n}}$$

$$j,e=1,\ldots,4 \quad j \neq e$$

where $q_{k,p}$ is tabulated under the title "Distribution of the Studentized Range" and

k = the number of sample means,

p = the number of degrees of freedom associated with the error mean square.

For the previous data generated by the single factor computer simulation experiment the formula for 95% confidence intervals is given by,

$$(EJ_j - EJ_e) = (\bar{J}_j - \bar{J}_e) \pm q_{4,36}\sqrt{\frac{214.66}{10.0}}$$

$$= (\bar{J}_j - \bar{J}_e) \pm (3.81)\sqrt{\frac{214.66}{10}}$$

$$= (\bar{J}_j - \bar{J}_e) \pm 17.65 \quad .$$

Table 4.4 contains a difference between sample means for all six pairs of difference in the experiment. An asterisk (*) indicates that the particular difference exceeds the confidence allowance, 17.65.

Similar results are given for the feedback relationship $C_i = \alpha_1 Y_{i-1} + u_i$.

The null hypothesis is likewise rejected for this case.

Table 4.4. Difference of sample means for $C_i = \alpha_1 Y_i + u_i$

α_1	j\e	2	3	4
.63	1	-13.94	-1.29	23.00*
.67	2		12.65	36.94*
.72	3			24.29*
.76	4			

Table 4.5. Difference of sample means for $C_i = \alpha_1 Y_{i-1} + u_i$

α_1	j\e	2	3	4	5
.63	1	-22.99*	-26.88*	-25.67*	-4.43
.67	2		-3.89	-2.68	18.56
.69	3			1.21	22.45*
.71	4				21.24*
.75	5				

Table 4.5 summarizes the results for all pairs of sample means, where the confidence interval is given by

$$(\bar{J}_j - \bar{J}_e) \pm q_{5,45}\sqrt{\frac{262.3}{10.}} \quad .$$

If the difference exceeds 20.58 it is significant at the .05 level and is indicated by the asterisk.

For the short time horizons of ten years in the non-linear problem T-3 the linear feedback relation was adequate to give results that agreed with the data computed from the control problem algorithms.

For a second feedback simulation experiment consider the T-3 problem with the following parameters. Choose B = .25 and set the rate of technical progress z equal to .03 to compute a model with parameter specifications approximating those of a developed economy. Let the time horizon be [0,50] and the other values are as given n = .9, δ = .05, r = .025, and the penalty constant pc = .1. The final stock of capital is chosen as K_T = 250. The results of deterministic simulation are summarized in Table 4.6.

Table 4.6. Objective function values for different choices of the feedback relation $C_i = \alpha_1 Y_i$ for time horizon [0,50] with B = .25, K_T = 250., z = .03, pc = .1 and a = 2.1723

J	α_1
96.0	.63
223.2	.65
299.4	.67
327.5	.69
310.0	.71
250.2	.73
150.8	.75

A regression of the optimal data computed from the control algorithm results in $C_i = .65\ Y_i$ with a coefficient of multiple determination of .91. The simulated optimization indicates the optimizing α_1 value as $\alpha_1 = .69$.

This value is close to the feedback value obtained by running the regression of C_i against Y_i from the time path data computed from the control algorithms. The stochastic simulation for the 50 year horizon and feedback relationship $C_i = \alpha_1 Y_i + u_i$ where u_i are normally independent and identically distributed with mean 0. and variance 1.0 follow in Table 4.7.

The differences between sample means are summarized in Table 4.8 where the asterisk indicates that the difference is significant at the .05 level (greater than 23.73).

Other feedback relationships could be considered as well as other parameter variations to solve the control problem by simulation, but these cases illustrate the feasibility of this alternative way of approximating the optimal solution to the control problem by simulation techniques.

Table 4.7. Average values of the objective function for different α_1 values where $C_i = \alpha_1 Y_i + u_i$ $u_i \approx$ N.I.D. (0,1.0)

Simulation run	$\overline{J}$	α_1
1	108.76	.63
2	292.31	.67
3	325.36	.69
4	309.42	.71
5	140.88	.75

Table 4.8. Difference in sample means

	2	3	4	5
1	-183.55*	-216.6*	-200.6*	-32.0*
2		-33.0	-17.11	151.5*
3			15.94	184.5*
4				168.5*

V. SUMMARY AND FURTHER RESEARCH

Four types of computation applied to optimal economic growth models have been considered and studied. The first was applying conjugate direction control algorithms to numerically solve deterministic optimal economic growth problems. The two iterative methods treated were the conjugate gradient and the Davidon algorithms and in both cases penalty functions were used to handle the terminal constraints on the state variables. In every case considered the Davidon method converged in less iterations and was less sensitive to the search direction parameter than the conjugate gradient method. The penalty function approach proved adequate to handle the terminal state constraints in all the problems that were studied. However a certain amount of numerical experimentation was needed to select the right magnitude for the penalty constants.

Experience with each problem was needed to determine the correct choices. Sequential unconstrainted minimization techniques of varying the penalty constants was helpful yet experimentation was still necessary to achieve good selections of the constants for each subproblem. Disadvantages of the conjugate direction iterative methods may be noted to include the difficulties encountered in treating inequality constraints. This requires penalty constants for the inequalities and for the terminal constraints. Success of the method is thus greatly dependent upon judicious choices of the penalty constants and requires a great deal of computer time and patience on the user's part to select adequate penalty constants.

However for nonlinear aggregative optimal growth models of from one to four state variables and one or possibly two control variables with no inequality and only terminal state constraints, the iterative conjugate direction algorithms appear from the computational experience reported earlier to be a reasonable choice to solve such nonlinear problems.

For operational planning type models with a large number of linear inequality constraints the second type of computational approach considered, the linear programming primal-dual problem with the neighboring extremal approach of Section A of Chapter 4 would be an attractive alternative to the iterative conjugate direction methods. A problem of further study would be to generalize the linear programming approach to a nonlinear objective functional.

The third computational approach considered in Section C of Chapter 3 was reducing a complex model to a less complex one by choosing some of the decision variables via a suboptimization procedure. This reduced the size of the problem and allowed for repeated solution of the complete model and the submodel by control theory iterative algorithms. This approach also allowed deterministic simulation on certain decision variables to approximate the optimal solutions.

The objective of the fourth computational approach, simulated optimization, is to locate improved but not necessarily optimum solutions. This technique is highly attractive compared to the computational difficulties involved in using iterative conjugate algorithms for large problems. The simulation can be utilized as described in Section C of Chapter 3 or as reported in Section B of Chapter 4. The latter approach proved successful in my experience reported in Chapter 4 Section B of assuming feedback relationships and optimizing on the parameters involved in the feedback relationship. The data generated can then be analyzed by a factorial experimental design and a comparison of cell means for the different choices of the parameters can then be made if the differences are significant. In addition, complex optimal economic growth models can also be reduced to simpler ones by assuming relationships between the system dynamics and/or state and control variables. The simple models can then be solved by iterative conjugate direction techniques for given functional relationships as were described in Section C of Chapter 3.

If the functional relationships to simplify the model are assumed stochastic, then various computed replications can be made where each modification is considered as a factor level in a factorial design. Significant differences and comparisons can then be made to approximate the optimal choice of the functional relationships. If deterministic relationships are assumed then deterministic simulation is effected. The advantage of this approach is that large optimal economic growth model solutions can be approximated by reducing the problem to a simpler one that can be solved by the iterative conjugate direction algorithms reported here. The solutions are computed for the various functional relationships and the statistical analysis performed. This eliminates the computational difficulties of a large control problem yet may increase the computer time (since each run is replicated) and sacrifices optimality for only an improved solution.

The feedback technique with a discretized model as reported in Section B of Chapter 4 is especially easy to compute. It requires no iterative control algorithms and uses only feedback relationships between the state and control variables. As all other simulated optimization approaches it only approximates the optimal solution. However, the simulation using feedback relationships allows for the easy incorporation of stochastic relationships in the model. This allows for more realism since the nature of many economic models would tend to be stochastic to allow for unpredictable factors rather than deterministic.

The feedback technique as presented in Section B of Chapter 4 certainly is an attractive procedure to solve either a deterministic or stochastic control problem if some idea of the state-control functional relationships are known.

Three areas for further research would include investigation into improved methods to handle inequality constraints in the control problem, continued investigation into the computation of stochastic control models as applied to economic growth, and investigation into the introduction of a nonlinear objective functional in the primal-dual linear programming approach of Section A, Chapter 4.

It is my plan to continue research activity dealing in these and related optimization areas.

VI. BIBLIOGRAPHY

1. Adachi, N., K. Goto and H. Tokumaru. Davidon's method for minimization problems in Hilbert space with an application to control problems. SIAM Journal on Control 8: 163-178. 1970.

2. Arrow, K. J. and L. Hurwicz. Decentralization and computation in resource allocation. In R. W. Pfouts, ed. Essays in economics and econometrics. Pp. 34-104. Chapel Hill, North Carolina, University of North Carolina Press. 1960.

3. Arrow, K. J. and M. Kurz. Public investment, the rate of return, and optimal fiscal policy. Baltimore, Maryland, The Johns Hopkins Press. c1970.

4. Athans, M. and P. L. Falb. Optimal control and introduction to the theory and its applications. New York, New York, McGraw-Hill Book Company, Inc. 1966.

5. Beckman, F. S. Solution of linear equations by the conjugate gradient method. In A. Ralston and H. S. Wilf, eds. Mathematical methods for digital computers. Pp. 62-72. New York, New York, John Wiley and Sons, Inc. 1960.

6. Bellman, R. Dynamic programming. Princeton, New Jersey, Princeton University Press. 1957.

7. Bjerve, P. J. Planning in Norway, 1947-1956. Amsterdam, Holland, North Holland Publishing Company. 1959.

8. Breakwell, J. V., J. L. Speyer and A. E. Bryson, Jr. Optimization and control of non-linear systems using the second variation. SIAM Journal on Control 1: 193-223. 1963.

9. Bruno, M. Optimal accumulation in discrete capital models. In K. Shell, ed. Essays in the theory of optimal economic growth. Pp. 181-218. Cambridge, Massachusetts, The M.I.T. Press. 1967.

10. Bryson, A. E., Jr. and W. F. Denham. A steepest-ascent method for solving optimum programming problems. Journal of Applied Mechanics 29: 247-252. 1962.

11. Bryson, A. E., Jr. and W. F. Denham. Optimum programming problems with inequality constraints: II. Solution by steepest ascent. AIAA Journal 2: 25-34. 1964.

12. Bryson, A. E., Jr. and Y. Ho. Applied optimal control. Waltham, Massachusetts, Blaisdell Publishing Company. c1969.

13. Burmeister, E. and A. R. Dobell. Mathematical theories of economic growth. New York, New York, Macmillan Company. c1970.

14. Central Planning Bureau of the Netherlands. Central economic plan. The Hague, Holland, author. 1961.

15. Central Planning Bureau of the Netherlands. Scope and methods of the central planning bureau. An exploration of economic potentialities of the Netherlands: 1950-1970. The Hague, Holland, author. 1956.

16. Chakravarty, S. Optimal saving with finite planning horizon. International Economic Review 3: 338-355. 1962.

17. Chenery, H. B. and M. Bruno. Development alternatives in an open economy: The case of Israel. Economic Journal 72: 79-103. 1962.

18. Chenery, H. B. The use of interindustry analysis in developmental programming. In T. Barna, ed. Structural independence and economic development. Pp. 20-21. London, England, Macmillan and Company. 1963.

19. Curry, H. B. The method of steepest descent for non-linear minimization problems. Quarterly of Applied Mathematics 2: 258-262. 1944.

20. Davidon, W. C. Variable metric method for minimization. Argonne National Laboratory Report ANL 5990 [Argonne National Labs., Argonne, Illinois]. Revised, February, 1966.

21. Dobell, A. R. Optimization in models of economic growth. In Studies in optimization 1. Pp. 1-27. Philadelphia, Pennsylvania, SIAM. 1970.

22. Fiacco, A. V. and G. P. McCormick. Computational algorithm for the sequential unconstrained minimization technique for nonlinear programming. Management Science 10: 601-617. 1964.

23. Fiacco, A. V. and G. P. McCormick. Extensions of SUMT for nonlinear programming: Equality constraints and extrapolation. Management Science 12: 816-820. 1966.

24. Fiacco, A. V. and G. P. McCormick. Nonlinear programming: Sequential unconstrained minimization techniques. New York, New York, John Wiley and Sons, Inc. c1968.

25. Fletcher, R. and M. J. D. Powell. A rapidly convergent descent method for minimization. Computer Journal 6: 163-168. 1963.

26. Fletcher, R. and C. M. Reeves. Function minimization by conjugate gradients. Computer Journal 7: 149-154. 1964.

27. Fox, K. A., J. K. Sengupta and E. Thorbecke. The theory of quantitative economic policy with applications to economic growth and stabilization. Chicago, Illinois, Rand McNally and Company. 1966.

28. Fromm, G. The evaluation of economic policies. In T. Naylor, ed. The design of computer simulation experiments. Pp. 355-369. Durham, North Carolina, Duke University Press. 1969.

29. Fromm, G. and P. Taubman. Policy simulation with an econometric model. Washington, D.C., The Brookings Institution. 1968.

30. Goodwin, R. M. Optimal growth path for an underdeveloped economy. Economic Journal 71: 756-774. 1961.

31. Hayes, R. M. Iterative methods of solving linear problems in a Hilbert space. National Bureau of Standards Applied Mathematics Series 39: 71-104. 1954.

32. Hestenes, M. R. Calculus of variations and optimal control theory. New York, New York, John Wiley and Sons, Inc. 1966.

33. Hestenes, M. R. and E. Stiefel. Methods of conjugate gradient for solving linear systems. Journal Research National Bureau of Standards 49: 409-436. 1952.

34. Hunter, J. S. and T. H. Naylor. Experimental design. In T. Naylor, ed. The design of computer simulation experiments. Pp. 39-58. Durham, North Carolina, Duke University Press. 1969.

35. Ijiri, Y. Management goals and accounting for control. Chicago, Illinois, Rand McNally and Company. 1965.

36. Intriligator, M. D. Mathematical optimization and economic theory. Englewood Cliffs, New Jersey, Prentice Hall, Inc. c1971.

37. Johansen, L. A. Durability of capital and the rate of growth of national product. International Economic Review 2: 361-370. 1961.

38. Jorgenson, D. W. The development of a dual economy. Economic Journal 71: 309-334. 1961.

39. Kelley, H. J. Gradient theory of optimal flight paths. American Rocket Society Journal 30: 947-953. 1960.

40. Klein, L. A model of Japanese economic growth. Econometrica 29: 277-292. 1961.

41. Klein, L. Some theoretical issues in the measurement of capacity. Econometrica 28: 272-286. 1960.

42. Lange, O. Some observations on input-output analysis. Sankhya 18: 310-324. 1957.

43. Lange, O. The computer and the market, presented in socialism, capitalism and economic growth. In C. H. Feinstein, ed. Essays presented to Maurice Dobb. Pp. 158-161. Cambridge, England, Cambridge University Press. 1967.

44. Lasdon, L. S. Conjugate direction methods for optimal control. IEEE Transactions on Automatic Control, AC-15: 267-268. 1970.

45. Lasdon, L. S., A. D. Waren and S. K. Mitter. The conjugate gradient method for optimal control problems. IEEE Transactions on Automatic Control, AC-12: 132-138. 1967.

46. Lasdon, L. S., A. D. Waren and R. K. Rice. An interior penalty method for inequality constrained optimal control problems. IEEE Transactions on Automatic Control, AC-12: 538-548. 1967.

47. Luenberger, D. G. Optimization by vector space methods. New York, New York, John Wiley and Sons, Inc. c1969.

48. Mahalanobis, P. C. The approach of operational research to planning in India. Sankhya 16: 3-62. 1955.

49. March, J. G. and H. A. Simon. Organizations. New York, New York, John Wiley and Sons, Inc. 1958.

50. McGill, R. Optimal control inequality state constraints and the generalized Newton-Raphson algorithm. SIAM Journal on Control 3: 291-297. 1965.

51. Mehra, R. K. and A. E. J. Bryson. Conjugate gradient methods with an application to V/STOL flight-path optimation. J. Aircraft 6: 123-128. 1969.

52. Morishima, M. Theory of economic growth. Oxford, England, Clarendon Press. 1969.

53. Narendra, K. S. and S. S. Tripathi. Optimization using conjugate gradient methods. IEEE Transactions on Automatic Control, AC-15: 268-270. 1970.

54. Naylor, T., J. Balintfy, D. Burdick and K. Chu. Computer simulation techniques. New York, New York, John Wiley and Sons Inc. c1966.

55. Naylor, T., K. Wertz and T. Wonnacott. Some methods for evaluating the effects of economic policies using simulation experiments. Review of the International Statistical Institute 36: 184-200. 1968.

56. Pearson, J. D. Variable metric methods of minimization. Computer Journal 12: 171-178. 1969.

57. Pierson, B. L. and S. W. Rajtora. IEEE Transactions on Systems Science and Cybernetics 6: 240-242. 1970.

58. Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko. The mathematical theory of optimal processes. New York, New York, Interscience Publishers. c1962.

59. Powell, M. J. D. A survey of numerical methods for unconstrained optimization. In Studies in optimization 1. Pp. 43-61. Philadelphia, Pennsylvania, SIAM. 1970.

60. Ramsey, F. P. A mathematical theory of saving. Economic Journal 38: 543-559. 1928.

61. Samuelson, P. A. A catenery turnpike theorem involving consumption and the golden rule. American Economic Review 55: 486-496. 1965.

62. Scheffe, H. The analysis of variance. New York, New York, John Wiley and Sons, Inc. c1959.

63. Sen, A. K. Terminal capital and optimum savings. Presented in socialism, capitalism and economic growth. In C. H. Feinstein, ed. Essays presented to Maurice Dobb. Pp. 40-53. Cambridge, Massachusetts, Cambridge University Press. 1967.

64. Sengupta, J. K. Some observations on the optimal growth path for an underdeveloped economy. Metroeconomica 20: 149-172. 1968.

65. Sengupta, J. K. and K. A. Fox. Optimization techniques in quantitative economic models. Amsterdam, Holland, North Holland Publishing Company. 1969.

66. Sengupta, J. K. and D. Walker. On the empirical specification of optimal economic policy for growth and stabilization under a macro-dynamic model, Manchester School of Economics and Social Science 32: 215-238. September, 1964.

67. Shell, K. Essays on the theory of optimal economic growth. Cambridge, Massachusetts, M.I.T. Press. c1967.

68. Sinnott, J. F., Jr. and D. G. Luenberger. Solution of optimal control problems by the method of conjugate gradients. Joint Automatic Control Conference Proc. 1967: 566-574. 1967.

69. Theil, H. Optimal decisions rules for government and industry. Chicago, Illinois, Rand McNally and Company. 1964.

70. Tripathi, S. S. and K. S. Narendra. Conjugate direction methods for nonlinear optimization problems. Proc. National Electronics Conference 1968: 125-129. December, 1968.

71. Willoughby, J. K. Adaptions of the conjugate gradient method to optimal control problems with terminal constraints. Contract no. N000-14-68A-0162. Washington D.C., Office of the Naval Research, Technical Report ER1-62500. Department of the Navy. February, 1970.

72. Uzawa, H. An optimal fiscal policy in an aggregative model of economic growth. In I. Adelman and E. Thorbecke, eds. The theory and design of economic development. Pp. 113-139. Baltimore, Maryland, The Johns Hopkins Press. 1966.

73. Vanden Bogaard, P. J. M. and H. Theil. Macro-dynamic policy-making: An application of strategy and certainty equivalence concepts to the economy of the United States 1933-36. Metroeconomica 11: 149-167. 1959.

VII. ACKNOWLEDGMENTS

Appreciation is extended to Dr. Jati K. Sengupta for his help and suggestions in the preparation of this dissertation.

VIII. APPENDIX A

A computer code for solving problem T-3 by the Davidon algorithm and penalty functions to treat terminal state constraints.

```
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON F,OPTSTP,PC,PSI,GSQ,DFA,T,TI,X(10),D(10),STOX1(101),STOX2(1
     101),STOU(101),G(101),S(101),GG(101),Y(101),IN,IEND,ITAB,IX,NFEVAL,
     2NREINT,NV,KOUN
      COMMON/MF/ PZ1,PZ2,PZ3,PZ4,PZ5,PZ6,PZ7,PZ8,PZ10
      PZ1=0.9
      PZ2=0.6
      PZ3=0.03
      PZ4=0.05
      PZ5=0.01
      PZ6=0.025
      PZ7=PZ6*(1.D0-PZ2)+PZ5
      PZ8=0.8419
      PZ10=1.D0/(1.D0-PZ1)
      IEND = 100
      ITAB = IEND + 1
       KOUN = 4
      PC=3.D0
      IMAX = 14
      TI =.1D0
      NV= 1
      ISTEP = 0
      IX = 1
      NFEVAL = 0
      NREINT = 0
      STOX1(1) = 15.D0
      STOU(1)=2.25D0
      T2Z = TI
      DO 10 I = 2,101
      STOU(I)=2.25D0+.35D0*T2Z
      T2Z = T2Z + TI
   10 S(I) = 0.D0
```

```
      CALL FANDG(0.DO)
      GSQ = SP(G,G,2)
      WRITE (6,100) (I,STOX1(I),G(I), I= 1,101,4)
100    FORMAT (I10,2D20.8)
      STPEST =  .1DO/DSQRT(GSQ)
    1 CONTINUE
      DO 2 I=1,ITAB
    2 S(I) = -G(I)
      DFA =-GSQ
    5 CONTINUE
      DO 19 I=1,ITAB
   19 GG(I) = G(I)
      CALL LINMIN(STPEST)
      ISTEP = ISTEP + 1
      IX = IX + 1
      DO 3 I=1,ITAB
    3 STOU(I) = STOU(I) + OPTSTP*S(I)
      WRITE (6,11) ISTEP,NFEVAL,NREINT,F,GSQ,PSI,(I,STOX1(I),STOX2(I),
     1             STOU(I),G(I),I=1,101,4)
       WRITE (6,17) OPTSTP
   17      FORMAT (/// D15.7//////)
      IF(ISTEP.EQ.4) GO TO 50
      IF(ISTEP.EQ.8) GO TO 50
      IF ( GSQ .LE. 1.D-04) GO TO 30
      IF(ISTEP.EQ.IMAX) GO TO 30
    3 STPEST = 4.DO*OPTSTP
      IF(STPEST.GT.1.DO) STPEST = 1.DO
      CALL SEAR
      IF (IX.EQ.1) GO TO 21
      DFA= SP(G,S,1)
      IF (DFA.GT.0.DO)   GO TO 6
      GO TO 5
    6 WRITE (6,12)
      IX = 1
      STPEST =  OPTSTP/10.DO
      GO TO 1
```

```
   21 STPEST =  OPTSTP/10.D0
      GO TO 1
   30 WRITE(6,40)
   40 FORMAT(1X,' TIME          OUTPUT       SAVINGS')
      T=0.D0
      DO 35 I=1,ITAB,4
      Y1=DEXP(PZ7*T)*PZ8*(STOX1(I))**PZ2
      S1=(Y1-STOU(I))/Y1
      WRITE(6,41) T,Y1,S1
   41 FORMAT(1H ,3F12.6)
   35  T = T + 4.D0*TI
   50 WRITE(6,51)
   51 FORMAT(1X,'    TIME ',20X,'HAMILTONIAN')
      TY1=0.D0
      DO 60 I=1,ITAB,10
      H1= -DEXP(-PZ3*TY1)*PZ10*(STOU(I))**(1.D0-PZ1)
      H2= STOX2(I)*(PZ8*DEXP(PZ7*TY1)*(STOX1(I))**PZ2-STOU(I)-PZ4*STOX1(
     1I))
      H=H1+H2
      WRITE(6,55) TY1,H
   55 FORMAT(1X,F12.6,D36.10)
   60 TY1=TY1+10.D0*TI
      IF (GSQ .LE. 1.D-04) STOP
      IF (ISTEP .EQ. IMAX ) STOP
      GO TO 18
   11 FORMAT (///1H0,8HSTEP.NO.,5X,20HFUNCTION EVALUATIONS,5X,16HREINTER
     1POLATIONS,20X,1HF,20X,3HGSQ,20X,3HPSI/1H0,I5,I18,I23,D36.10,D20.6,
     2D23.6/1H0,32X,5HINDEX,13X,2HX1,18X,2HX2,19X,1HU,19X,1HG//(1H ,I36,
     33X,4D20.8))
   12 FORMAT ('0UPHILL DIRECTION OF SEARCH--A STEEPEST DESCENT STEP WILL
     1FOLLOW')
      END
```

```
      SUBROUTINE FANDG(TSTEP)
      IMPLICIT REAL*8(A-H,C-Z)
      COMMON F,OPTSTP,PC,PSI,GSQ,DFA,T,TI,X(10),D(10),STOX1(101),STOX2(1
     1101),STOU(101),G(101),S(101),GG(101),Y(101),IN,IEND,ITAB,IX,NFEVAL,
     2NREINT,NV,KOUN
      COMMON/MF/ PZ1,PZ2,PZ3,PZ4,PZ5,PZ6,PZ7,PZ8,PZ10
      DIMENSION Z(101)
C     SAVE THE STORED CONTROL TABLE BY TRANSFERRING ITS CONTENTS TO Z
C     TABLE Z IS USED AS THE CONTROL IN THIS SUBROUTINE
C
      DO 1 I=1,ITAB
    1 Z(I) = STOU(I) + TSTEP*S(I)
C
C     INTEGRATE THE  STATE SYSTEM
C
      T = .0D0
      X(1) = 15.D0
      I=1
      IN= 1
    2 IF(IN.EQ.4 .OR. IN.EQ.13)    GO TO 3
      IF(IN.EQ.3) U=(Z(I)+ Z(I+1))/2.D0
      IF(IN.EQ.5) U= Z(I+1)
      IF(IN.EQ.1) U= Z(1)
    3 X1 =X(1)
      D(1)=(PZ8*DEXP(PZ7*T)*(X1)**PZ2-U-PZ4*X1)
    4 CALL STEP(&2)
      I=I+1
      STOX1(I) = X(1)
      IF(I .LT. ITAB) GO TO 4
      PSI = X(1) - 29.16D0
      PCPSI =(PC)*PSI
      F=(PZ10*(Z(1)**(1.DC-PZ1)))/2.D0
```

```
      F = - F
      TR = 0.D0
      DO 40 I=2,IEND
      TR = TR + TI
   40 F=F-(DEXP(-PZ3*TR))*PZ10*(Z(I)**(1.D0-PZ1))
      TR=TR+TI
      F=F-(DEXP(-PZ3*TR)*PZ10*(Z(ITAB)**(1.D0-PZ1)))/2.D0
C
C     INTEGRATE ADJOINT  EQUATIONS
      F = TI*F + .5D0* PCPSI*PSI
C     CALCULATE THE  GRADIENT OF THE  HAMILTONIAN
C
      T = 10.D0
      TI = -0.1D0
      X(1) =PCPSI
      STOX2(ITAB) = PCPSI
      G(ITAB)=-DEXP(-PZ3*T)*(Z(ITAB)**(-PZ1))-X(1)
      I = ITAB
      IN = 1
    5 IF(IN.EQ.4 .OR. IN.EQ.13)   GO TO 6
      IF( IN.EQ.3)   GO TO 10
      IF( IN.EQ.5)   GO TO 20
      IF( IN.EQ.1)   GO TO 30
      GO TO 12
   30 X1 = STOX1(ITAB)
      U= Z(ITAB)
      GO TO 12
   20 X1 = STOX1(I-1)
      U= Z(I-1)
      GO TO 12
   10 X1 = (STOX1(I) + STOX1(I-1))/2.D0
      U=(Z(I) + Z(I-1))/2.D0
   12  CONTINUE
  6   D(1)=X(1)*(PZ4-(PZ8*PZ2*(DEXP(PZ7*T))*X1**(PZ2-1.D0)))
    7 CALL STEP(&5)
      I=I-1
```

```
      STOX2(I) = X(1)
      G(I)=-DEXP(-PZ3*T)*(Z(I)**(-PZ1))-X(1)
      IF(I.GT.1)   GO TO 7
      TI =0.1D0
      NFEVAL = NFEVAL + 1
      RETURN
      END

      SUBROUTINE STEP(*)
      IMPLICIT REAL*8(A-H,O-Z), INTEGER (I-N)
      COMMON F,OPTSTP,PC,PSI,GSC,DFA,T,TI,X(10),D(10),STOX1(101),STOX2(1
     101),STOU(101),G(101),S(101),GG(101),Y(101),IN,IEND,ITAB,IX,NFEVAL,
     2NREINT,NV,KOUN
      DIMENSION XS(10),DS(10),Z(10),XD(10)
      GO TO (10,50,1020,1040,1060,50,50,50,50,50,50,50,1320,5000),IN
   10 TD =T
      TS =T
      DO 20 I=1,NV
           XD(I) = X(I)
   20      DS(I) = D(I)
 5000 H=TI
 1000 H2 = .5D0*H
      H6 = H2/3.D0
      T=TS + H2
      DO 1010 I=1,NV
          XS(I) = XD(I)
 1010     X(I)  = XS(I) +H2*DS(I)
      IN = 3
      RETURN 1
C
 1020 DO 1030 I=1,NV
        DD = D(I)
      Z(I) = DS(I) + 2.D0*DD
 1030 X(I) = XS(I) + H2*DD
      IN = 4
      RETURN 1
```

```
1040 T= T +H2
     DO 1050 I=1,NV
        DD = D(I)
        Z(I) = Z(I) + 2.D0*DD
1050    X(I) = XS(I) + H*DD
     IN = 5
     RETURN 1
1060 DO 1070 I=1,NV
1070        Z(I) = H6*(Z(I)+D(I))
1300 TD=TD +H
        TS = TD
     T = TS
     DO 1310 I=1,NV
         XD(I) = XD(I) + Z(I)
1310     X(I) = XD(I)
     IN = 13
     RETURN 1
1320 DO 1330 I=1,NV
1330      DS(I) = D(I)
  50 CONTINUE
5020 IN = 14
     RETURN
     END

      FUNCTION SP(XX,YY,IMODE)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON F,CPTSTP,PC,PSI,GSQ,DFA,T,TI,X(10),D(10),STOX1(101),STOX2(1
     101),STCU(101),G(101),S(101),GG(101),Y(101),IN,IEND,ITAB,IX,NFEVAL,
     2NREINT,NV,KCUN
      DIMENSION XX(101),YY(101)
      IF(IMODE.EQ.1)  GO TO 1
      IF(IMODE.EQ.2)  GO TO 2
      SP = XX(1)/2.0D0
      DO 3 I=2,IEND
    3 SP = SP + XX(I)
      SP = SP + XX(ITAB)/2.D0
```

```
      SP = TI*SP
      RETURN
    1 SP = XX(1)*YY(1)/2.D0
      DO 4 I=2,IEND
    4 SP = SP + XX(I)*YY(I)
      SP = SP + XX(ITAB)*YY(ITAB)/2.D0
      SP = TI*SP
      RETURN
    2 SP = XX(1)*XX(1)/2.D0
      DO 5 I=2,IEND
      Z = XX(I)
    5 SP = SP + Z*Z
      SP = SP + XX(ITAB)*XX(ITAB)/2.D0
      SP = TI*SP
       RETURN
      END

      SUBROUTINE LINMIN(STPEST)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON STOX1(101),STOX2(101),STCU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,DFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IEND,ITAB
  342 FORMAT('0ALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'OPTSTP=',D14.6)
  343 FORMAT(' ALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'STPEST=',D14.6)
  344 FORMAT(' ALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'OPTSTP=',D14.6)
      IWORK=0
      SSQ = SP(S,S,2)
      ALPHA=0.D0
  300 FA=F
  301 BETA=ALPHA+STPEST
      CALL FANDG(BETA)
      DFB = SP(G,S,1)
      WRITE(6,343)ALPHA,BETA,DFA,DFB,F,STPEST
```

```
      IF ( DFB .GT. 1.D18 ) GC TO 310
      IF ( F.GT.FA .OR. DFB.GT. 0.D0 )  GO TO 302
      ALPHA = BETA
      DFA = DFB
      STPEST = 4.D0*STPEST
      GO TO 300
  302 OLDF = F
      FB=F
  303 U = DFA + DFB + 3.D0*( (FA-FB) / (BETA-ALPHA) )
      W = DSQRT( U*U - DFA*DFB )
      FACTOR = ( DFB + W - U ) / ( DFB - DFA + 2.D0*W )
      IF ( FACTOR.GE.1.D0 .OR. FACTOR.LT. 0.D0 )  GO TO 311
      OPTSTP = BETA - FACTCR*( BETA - ALPHA )
  330 CALL FANDG(OPTSTP)
      GSQ = SP(G,G,2)
      WRITE(6,342)ALPHA,BETA,DFA,DFB,F,OPTSTP
      DFOPT = SP(G,S,1)
      IF ( F.GT.FA .OR. F.GT.FB )   GC TC 399
      IF ( DFOPT*DFOPT/(GSC*SSQ) .LT. C.0004D0 )   RETURN
  399 CONTINLE
      IF ( IWORK.GE.5 .AND. DABS(OLDF-F).LE.1.D-07 )  GO TO 306
      NN = 0
      OLDF=F
      IWORK = IWORK + 1
      NREINT = NREINT + 1
      IF ( F.GT.FA .OR. DFOPT.GT.C.D0)  GO TO 312
      IF ( OPTSTP .GT. .7D0*ALPHA+.3D0*BETA )   NN=1
      ALPHA = OPTSTP
      FA = F
      DFA = DFUPT
      IF (NN.EQ.1)   GC TO 303
      TSTEP = 0.5D0*(ALPHA+BETA)
      CALL FANDG(TSTEP)
      WRITE(6,344)ALPHA,BETA,CFA,DFB,F,OPTSTP
      DF = SP(G,S,1)
      IF ( F.GT.FA .OR. DF.GT.0.DC)  GC TO 320
```

```
      GO TO 303
  306 WRITE (6,341)
      RETURN
  310 STPEST = 0.1D0 * STPEST
      GO TO 301
  311 WRITE( 6, 340 ) FACTOR
      OPTSTP = .5D0 * ( ALPHA + BETA )
      GO TO 330
  312 IF ( OPTSTP .LT. .3D0*ALPHA+.7D0*BETA )   NN=1
      BETA = OPTSTP
      FB = F
      DFB = DFOPT
      IF (NN.EQ.1)   GO TO 303
      TSTEP = 0.5D0*(ALPHA+BETA)
      CALL FANDG(TSTEP)
      WRITE(6,344)ALPHA,BETA,DFA,DFB,F,OPTSTP
      DF = SP(G,S,1)
      IF ( F.LE.FA .AND. DF.LT.0.D0 )   GO TO 321
      GO TO 303
  320 DFB = DF
      FB = F
      BETA = TSTEP
      GO TO 303
  321 DFA = DF
      ALPHA = TSTEP
      FA = F
      GO TO 303
  340 FORMAT(' UNACCEPTABLE FACTOR CHOSEN     FACTOR = ',D12.5)
  341 FORMAT (1H0,32HLAST RESORT EXIT TAKEN IN LINMIN)
      END
```

```
      SUBROUTINE SEAR
      IMPLICIT REAL*8(A-H,C-Z)
      COMMON F,OPTSTP,PC,PSI,GSG,DFA,T,TI,X(10),D(10),STOX1(101),STOX2(1
     101),STOU(101),G(101),S(101),GG(101),Y(101),IN,IEND,ITAB,IX,NFEVAL,
     2NREINT,NV,KOUN
      DIMENSION SIG(101),HY(101),XS( 7,101),XSS( 7,101)
      EQUIVALENCE(SIG(1),HY(1))
       IF ( IX .NE. KOUN) GO TO 1
      IX=1
      WRITE(6,100)
  100 FORMAT('1SYSTEM HAS BEEN RESTARTED')
      RETURN
    1 DO 2 I=1,ITAB
      Y(I)= G(I)-GG(I)
    2 SIG(I)= OPTSTP*S(I)
      II=IX-1
      A= SP(SIG,Y,1)**(-0.5D0)
      DO 3 I=1,ITAB
      XS(II,I)= A*SIG(I)
    3 HY(I)=Y(I)
      IF(II.EQ.1) GO TO 12
      III=II-1
      DO 4 N=1,III
C  SWITCH XS(N,I) TO S(I) AND XSS(N,I) TO GG(I) SINCE THEY ARE NO LONGER NE
      DO 5 I=1,ITAB
      S(I)=XS(N,I)
    5 GG(I)=XSS(N,I)
      A= SP(S,Y,1)
      B= SP(GG,Y,1)
      DO 6 I=1,ITAB
    6 HY(I)=HY(I)+ A*S(I)- B*GG(I)
    4 CONTINUE
   12 CONTINUE
      A= SP(HY,Y,1)**(-0.5D0)
      DO 7 I=1,ITAB
    7 XSS(II,I)=A*HY(I)
```

```
      DO 8 I=1,ITAB
    8 S(I)=-G(I)
      DO 9 N=1,II
C  SWITCH XS(N,I) TO Y(I) AND XSS(N,I) TO HY(I) SINCE THEY ARE NOT NEEDED
      DO 10 I=1,ITAB
      Y(I)= XS(N,I)
   10 HY(I)=XSS(N,I)
      A= SP(Y,G,I)
      B= SP(HY,G,I)
      DO 11 I=1,ITAB
   11 S(I)= S(I)- A*Y(I)+ B*HY(I)
    9 CONTINUE
      RETURN
      END
```

IX. APPENDIX B

A computer code for solving problem T-3 using the conjugate gradient algorithms and penalty functions to treat terminal state constraints.

```
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON STOX1(101),STOX2(101),STOU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,DFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IEND,ITAB
      COMMON/MF/ PZ1,PZ2,PZ3,PZ4,PZ5,PZ6,PZ7,PZ8,PZ10
      PZ1=0.9
      PZ2=0.6
      PZ3=0.03
      PZ4=0.05
      PZ5=0.01
      PZ6=0.025
      PZ7=PZ6*(1.D0-PZ2)+PZ5
      PZ8=0.8419
      PZ10=1.D0/(1.D0-PZ1)
      PC=3.D0
      IEND = 100
      ITAB = IEND + 1
      IMAX = 16
      TI =.1D0
      NV= 1
      ISTEP = 0
      NFEVAL = 0
      NREINT = 0
      STOX1(1) = 15.D0
      STOU(1) = 2.25D0
      T2Z = TI
      DO 10 I = 2,101
      STOU(I) = 2.25D0 + .35D0*T2Z
      T2Z = T2Z + TI
   10 S(I) = 0.D0
      CALL FANDG(0.D0)
      GSQ = SP(G,G,2)
      STPEST =  .1D0/DSQRT(GSQ)
    1 CONTINUE
      DO 2 I=1,ITAB
    2 S(I) = -G(I)
      DFA =-GSQ
```

```
    5 CONTINUE
      OLDGS = GSQ
      CALL LINMIN(STPEST)
      ISTEP = ISTEP + 1
      DO 3 I=1,ITAB
    3 STOU(I) = STOU(I) + OPTSTP*S(I)
      GSQ = SP(G,G,2)
      WRITE (6,11) ISTEP,NFEVAL,NREINT,F,GSQ,PSI,(I,STOX1(I),STOX2(I),
     1                STOU(I),G(I),I=1,101,4)
       WRITE (6,17) OPTSTP
   17     FORMAT (/// D15.7//////)
      IF(ISTEP.EQ.4) GO TO 50
      IF(ISTEP.EQ.8) GO TO 50
      IF ( GSQ .LE. 1.D-04) GO TO 30
      IF(ISTEP.EQ.IMAX) GO TC 30
   18 STPEST = 2.D0*OPTSTP
       IF ( STPEST .GT. 1.D0)  STPEST = .5D0
      BETA = GSQ/OLDGS
      DO 4 I=1,ITAB
    4 S(I) = -G(I) + BETA*S(I)
      DFA= SP(G,S,1)
      IF (DFA.GT.0.D0)   GO TC 6
      GO TO 5
    6 WRITE (6,12)
      GO TO 1
   30 WRITE(6,40)
   40 FORMAT(1X,' TIME        CUTPUT       SAVINGS')
      T=0.D0
      DO 35 I=1,ITAB,4
      Y1=DEXP(PZ7*T)*PZ8*(STCX1(I))**PZ2
      S1=(Y1-STOU(I))/Y1
      WRITE(6,41) T,Y1,S1
   41 FORMAT(1H ,3F12.6)
   35  T = T + 4.DC*TI
   50 WRITE(6,51)
   51 FORMAT(1X,'   TIME ',20X,'HAMILTCNIAN')
```

```
      TY1=0.D0
      DO 60 I=1,ITAB,10
      H1= -DEXP(-PZ3*TY1)*PZ10*(STOU(I))**(1.D0-PZ1)
      H2= STOX2(I)*(PZ8*DEXP(PZ7*TY1)*(STOX1(I))**PZ2-STOU(I)-PZ4*STOX1(
     1I))
      H=H1+H2
      WRITE(6,55) TY1,H
   55 FORMAT(1X,F12.6,D36.10)
   60 TY1=TY1+10.D0*TI
      IF (GSQ .LE. 1.D-04) STOP
      IF (ISTEP .EQ. IMAX ) STOP
      GO TO 18
   11 FORMAT (///1H0,8HSTEP.NO.,5X,20HFUNCTION EVALUATIONS,5X,16HREINTER
     1POLATIONS,20X,1HF,20X,3HGSQ,20X,3HPSI/1H0,I5,I18,I23,D36.10,D20.6,
     2D23.6/1H0,32X,5HINDEX,13X,2HX1,18X,2HX2,19X,1HU,19X,1HG//(1H ,I36,
     33X,4D20.8))
   12 FORMAT ('CUPHILL DIRECTION OF SEARCH--A STEEPEST DESCENT STEP WILL
     1FOLLOW')
      END

      SUBROUTINE FANDG(TSTEP)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON STOX1(101),STOX2(101),STOU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,UFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IEND,ITAB
      COMMON/MF/ PZ1,PZ2,PZ3,PZ4,PZ5,PZ6,PZ7,PZ8,PZ10
      DIMENSION Z(101)
C     SAVE THE STORED CONTROL TABLE BY TRANSFERRING ITS CONTENTS TO Z
C     TABLE Z IS USED AS THE CONTROL IN THIS SUBROUTINE
C
      DO 1 I=1,ITAB
    1 Z(I) = STOU(I) + TSTEP*S(I)
C
C     INTEGRATE THE  STATE SYSTEM
C
      T= 0.D0
      X(1) = 15.D0
```

```
      I=1
      IN= 1
    2 IF(IN.EQ.4 .OR. IN.EQ.13)    GO TO 3
      IF(IN.EQ.3) U=(Z(I)+ Z(I+1))/2.DO
      IF(IN.EQ.5) U= Z(I+1)
      IF(IN.EQ.1) U= Z(1)
    3 XI =X(1)
      D(1)=(PZ8*DEXP(PZ7*T)*(X1)**PZ2-U-PZ4*X1)
    4 CALL STEP(&2)
      I=I+1
      STOX1(I) = X(1)
      IF(I .LT. ITAB) GO TO 4
      PSI = X(1) - 29.16DO
      PCPSI =(PC)*PSI
      F=(PZ10*(Z(1)**(1.DO-PZ1)))/2.DO
      F = - F
      TR = 0.DO
      DO 40 I=2,IEND
      TR = TR + TI
   40 F=F-(DEXP(-PZ3*TR))*PZ10*(Z(I)**(1.DO-PZ1))
      TR=TR+TI
      F=F-(DEXP(-PZ3*TR)*PZ10*(Z(ITAB)**(1.DO-PZ1)))/2.DO
      F = TI*F + .5DO* PCPSI*PSI
C
C     INTEGRATE ADJOINT  EQUATIONS
C     CALCULATE THE  GRADIENT OF THE  HAMILTONIAN
C
      T= 10.DO
      TI = -0.1DO
      X(1) =PCPSI
      STOX2(ITAB) = PCPSI
      G(ITAB)=-DEXP(-PZ3*T)*(Z(ITAB)**(-PZ1))-X(1)
      I = ITAB
      IN = 1
    5 IF(IN.EQ.4 .OR. IN.EQ.13)    GO TO 6
      IF( IN.EQ.3)   GO TO 10
```

```
      IF( IN.EQ.5)    GC TO 2C
      IF( IN.EQ.1)    GC TO 30
      GO TO 12
   30 X1 = STCX1(ITAB)
      U= Z(ITAB)
      GO TO 12
   2C X1 = STCX1(I-1)
      U= Z(I-1)
      GO TC 12
   1C X1 = (STOX1(I) + STOX1(I-1))/2.DO
      U=(Z(I) + Z(I-1))/2.DO
   12  CONTINUE
  6   D(1)=X(1)*(PZ4-(PZ8*PZ2*(DEXP(PZ7*T))*X1**(PZ2-1.DO)))
    7 CALL STEP(&5)
      I=I-1
      STOX2(I) = X(1)
      G(I)=-DEXP(-PZ3*T)*(Z(I)**(-PZ1))-X(1)
      IF(I.GT.1)    GO TO 7
      TI =0.1DO
      NFEVAL = NFEVAL + 1
      RETURN
      END

      SUBROUTINE STEP(*)
      IMPLICIT REAL*8(A-H,O-Z), INTEGER (I-N)
      COMMON STOX1(101),STCX2(101),STOU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,DFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IENC,ITAB
      DIMENSION XS(10),DS(10),Z(10),XD(1C)
      GO TO (10,50,102C,1040,106C,50,50,50,50,50,50,50,1320,5000),IN
   10 TD =T
      TS =T
      DO 20 I=1,NV
           XD(I) = X(I)
   20      DS(I) = D(I)
 5000 H=TI
```

```
 1000 H2 = .5D0*H
      H6 = H2/3.D0
      T=TS + H2
      DO 1010 I=1,NV
          XS(I) = XD(I)
 1010     X(I)  = XS(I) +H2*DS(I)
      IN = 3
      RETURN 1
C
 1020 DO 1030 I=1,NV
        DD = D(I)
      Z(I) = DS(I) + 2.D0*DD
 1030 X(I) = XS(I) + H2*DD
      IN = 4
      RETURN 1
 1040 T= T +H2
      DO 1050 I=1,NV
         DD = D(I)
         Z(I) = Z(I) + 2.D0*DD
 1050    X(I) = XS(I) + H*DD
      IN = 5
      RETURN 1
 1060 DO 1070 I=1,NV
 1070        Z(I) = H6*(Z(I)+D(I))
 1300 TD=TD +H
         TS = TD
      T = TS
      DO 1310 I=1,NV
          XD(I) = XD(I) + Z(I)
 1310     X(I) = XD(I)
      IN = 13
      RETURN 1
 1320 DO 1330 I=1,NV
 1330       DS(I) = D(I)
 5020 IN = 14
   50 CONTINUE
      RETURN
      END
```

```
      FUNCTION SP(XX,YY,IMODE)
      IMPLICIT REAL*8(A-H,O-Z)
      COMMON STOX1(101),STOX2(101),STCU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,DFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IEND,ITAB
      DIMENSION XX(101),YY(101)
      IF(IMODE.EQ.1)  GO TO 1
      IF(IMODE.EQ.2)  GO TO 2
      SP = XX(1)/2.0D0
      DO 3 I=2,IEND
    3 SP = SP + XX(I)
      SP = SP + XX(ITAB)/2.D0
      SP = TI*SP
      RETURN
    1 SP = XX(1)*YY(1)/2.D0
      DO 4 I=2,IEND
    4 SP = SP + XX(I)*YY(I)
      SP = SP + XX(ITAB)*YY(ITAB)/2.D0
      SP = TI*SP
      RETURN
    2 SP = XX(1)*XX(1)/2.D0
      DO 5 I=2,IEND
      Z = XX(I)
    5 SP = SP + Z*Z
      SP = SP + XX(ITAB)*XX(ITAB)/2.D0
      SP = TI*SP
       RETURN
      END
```

```
      SUBROUTINE LINMIN(STPEST)
      IMPLICIT REAL*8(A-H,C-Z)
      COMMON STOX1(101),STOX2(101),STCU(101),G(101),S(101),F,OPTSTP,PC,
     1PSI,GSQ,DFA,NFEVAL,NREINT,T,TI,X(10),D(10),NV,IN,IEND,ITAB
  342 FORMAT('OALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'UPTSTP=',D14.6)
  343 FORMAT(' ALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'STPEST=',D14.6)
  344 FORMAT(' ALPHA=',D14.6,2X,'BETA=',D14.6,2X,'DFA=',D14.6,2X,'DFB=',
     1D14.6,2X,'F=',D16.8,2X, 'OPTSTP=',D14.6)
      IWORK=0
      SSQ = SP(S,S,2)
      ALPHA=0.DO
  300 FA=F
  301 BETA=ALPHA+STPEST
      CALL FANDG(BETA)
      DFB = SP(G,S,1)
      WRITE(6,343)ALPHA,BETA,DFA,DFB,F,STPEST
      IF ( DFB .GT. 1.D18 ) GC TO 310
      IF ( F.GT.FA .OR. DFB.GT. 0.DO )  GO TO 302
      ALPHA = BETA
      DFA = DFB
      STPEST = 4.DO*STPEST
      GO TO 300
  302 OLDF = F
      FB=F
  303 U = DFA + DFB + 3.DO*( (FA-FB) / (BETA-ALPHA) )
      W = DSQRT( U*U - DFA*DFB )
      FACTOR = ( DFB + W - U ) / ( DFB - DFA + 2.DO*W )
      IF ( FACTCR.GE.1.DO .OR. FACTOR.LT. 0.DO )  GO TO 311
      OPTSTP = BETA - FACTCR*( BETA - ALPHA )
  330 CALL FANDG(OPTSTP)
      GSQ = SP(G,G,2)
      WRITE(6,342)ALPHA,BETA,DFA,DFB,F,OPTSTP
      DFOPT = SP(G,S,1)
      IF ( F.GT.FA .OR. F.GT.FB )   GC TC 399
```

```
      IF ( DFOPT*DFOPT/(GSQ*SSQ) .LT. 0.0004D0 )   RETURN
  399 CONTINUE
      IF ( IWORK.GE.5 .AND. DABS(OLDF-F).LE.1.D-07 )  GO TO 306
      NN = 0
      OLDF=F
      IWORK = IWORK + 1
      NREINT = NREINT + 1
      IF ( F.GT.FA .OR. DFOPT.GT.0.D0)  GO TO 312
      IF ( OPTSTP .GT. .7D0*ALPHA+.3D0*BETA )   NN=1
      ALPHA = OPTSTP
      FA = F
      DFA = DFOPT
      IF (NN.EQ.1)   GO TO 303
      TSTEP = 0.5D0*(ALPHA+BETA)
      CALL FANDG(TSTEP)
      WRITE(6,344)ALPHA,BETA,DFA,DFB,F,OPTSTP
      DF = SP(G,S,1)
      IF ( F.GT.FA .OR. DF.GT.0.D0)  GO TO 320
      GO TO 303
  306 WRITE (6,341)
      RETURN
  310 STPEST = 0.1D0 * STPEST
      GO TO 301
  311 WRITE( 6, 340 ) FACTOR
      OPTSTP = .5D0 * ( ALPHA + BETA )
      GO TO 330
  312 IF ( OPTSTP .LT. .3D0*ALPHA+.7D0*BETA )   NN=1
      BETA = OPTSTP
      FB = F
      DFB = DFOPT
      IF (NN.EQ.1)   GO TO 303
      TSTEP = 0.5D0*(ALPHA+BETA)
      CALL FANDG(TSTEP)
      WRITE(6,344)ALPHA,BETA,DFA,DFB,F,OPTSTP
      DF = SP(G,S,1)
      IF ( F.LE.FA .AND. DF.LT.0.D0 )   GO TO 321
```

```
      GO TO 303
  320 DFB = DF
      FB = F
      BETA = TSTEP
      GO TO 303
  321 DFA = DF
      ALPHA = TSTEP
      FA = F
      GO TO 303
  340 FORMAT(' UNACCEPTABLE FACTOR CHCSEN     FACTOR = ',D12.5)
  341 FORMAT (1HC,32HLAST RESCRT EXIT TAKEN IN LINMIN)
      END
```