OPTIMAL CONTROL FOR STERILIZATION
OF CANNED FOODS

by

Manoj Mangesh Nadkarni

B.Tech., Indian Institute of Technology, Bombay (1981)

Submitted to the Department of
Chemical Engineering
in Partial Fulfillment of the
Requirements of the Degree of

MASTER OF SCIENCE IN CHEMICAL ENGINEERING

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1983

© Massachusetts Institute of Technology 1983

Signature of Author: ____________________________

Department of Chemical Engineering
2 August, 1983

Certified by: ____________________________

T. Alan Hatton, Thesis Supervisor

Accepted by: ____________________________

Chairman, Departmental Graduate Committee

Archives

MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

OCT 24 1983

LIBRARIES
OPTIMAL CONTROL FOR STERILIZATION
OF CANNED FOODS

by

MANOJ MANGESH NADKARNI

Submitted to the Department of Chemical Engineering
on August 2, 1983 in partial fulfillment of the
requirements for the Degree of Master of Science in
Chemical Engineering

ABSTRACT

A mathematical model has been developed for maximizing the
nutrient retention during thermal sterilization of canned
foods. Pontryagin's minimum principle as applied to distrib-
uted parameter systems, has been used in the analysis. The
necessary conditions for optimization, derived by variational
methods, lead to bang-bang type of control which requires the
sterilization to be carried out at maximum rates of heating
and cooling.

A numerical procedure is outlined for the solution of the
resulting two-point boundary-value problem, which consists of
a system of nonlinear, coupled partial differential equations.
The unique behaviour of the costate variables makes it pos-
sible to integrate only the state equations in the forward
direction in the time domain, and to iterate on the final
desired reduction of micro organisms. A single point switching
shows less nutrient degradation than any combination of
multiple switchings.

Optimization with constraints on the retort temperature
results in a dual point switching where the control is either
at a maximum or minimum, or is zero. The model has been
applied to the sterilization of pork puree. Comparison with
the nutrient retention obtained using other temperature
policies, shows that bang-bang control with single point
switching represents the optimal solution.

Thesis Supervisor: Dr. T. Alan Hatton
Title: Assistant Professor of Chemical Engineering
ACKNOWLEDGEMENTS

This is a necessary, but not a sufficient expression of my indebtedness to several individuals. Undoubtedly, the person to whom I owe the most, is my thesis advisor, Prof. T. Alan Hatton, whose able guidance and endless efforts will never be forgotten. His continuous encouragement and the willingness to spend long hours of discussions have enabled me to complete this thesis.

Among others, I am grateful to Prof. Robert A. Brown for useful discussions on numerical methods, both in the class as well as outside the class. A word of very special thanks goes to Prof. Robert R. Tenney of the department of electrical engineering at MIT, for patiently explaining to me many intricacies of the optimal control theory. His help is greatly appreciated. The financial support during this research work was provided by MIT.

There are many friends at MIT and elsewhere, who have helped me in different ways, perhaps most importantly, by giving a moral support. I owe it to all of them. Finally, in a humble appreciation for what he has done for me, I wish to dedicate this thesis to my father.
<table>
<thead>
<tr>
<th>Number</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Results of Saguy and Karel (1979)</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>Modified Control for Constraint on the Retort Temperature</td>
<td>60</td>
</tr>
<tr>
<td>5.1</td>
<td>Basis for Obtaining a Single Point Switching</td>
<td>71</td>
</tr>
<tr>
<td>5.2</td>
<td>Condition for Optimal Switching</td>
<td>74</td>
</tr>
<tr>
<td>5.3</td>
<td>Condition for Switching for the case of no Constraint on the Final Concentration of Micro Organisms</td>
<td>75</td>
</tr>
<tr>
<td>5.4</td>
<td>Bang-Bang Control: Single Switching</td>
<td>77</td>
</tr>
<tr>
<td>5.5</td>
<td>Final Concentration of Micro Organisms as a Function of Switching Time</td>
<td>79</td>
</tr>
<tr>
<td>5.6</td>
<td>Results for Computer Simulation for Bang-Bang Control</td>
<td>80</td>
</tr>
<tr>
<td>5.7</td>
<td>Effect of Variation in the Limits of the Control u</td>
<td>81</td>
</tr>
<tr>
<td>5.8</td>
<td>Variation in Process Time</td>
<td>83</td>
</tr>
<tr>
<td>5.9</td>
<td>Nutrient Degradation at Different Activation Energies</td>
<td>84</td>
</tr>
<tr>
<td>5.10</td>
<td>Optimal Control for Upper Constraint on the Retort Temperature</td>
<td>86</td>
</tr>
<tr>
<td>5.11</td>
<td>Optimal Control with Lower Constraint on the Retort Temperature</td>
<td>87</td>
</tr>
<tr>
<td>5.12</td>
<td>Bang-Bang Control: Two Switching Points</td>
<td>94</td>
</tr>
<tr>
<td>5.13</td>
<td>Results for Two Switching Points</td>
<td>95</td>
</tr>
<tr>
<td>5.14</td>
<td>Optimal Concentration Using Two Switching Points</td>
<td>96</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.15</td>
<td>Optimal Control: Three Switching Points</td>
<td>97</td>
</tr>
<tr>
<td>5.16</td>
<td>Optimal Concentration with Three Switching Points</td>
<td>99</td>
</tr>
<tr>
<td>5.17</td>
<td>Variation in Second Switching Time $t_2$</td>
<td>99</td>
</tr>
<tr>
<td>(Three Switching Points)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1</td>
<td>Maximum Nutrient Retention at Constant Uniform Temperature</td>
<td>103</td>
</tr>
<tr>
<td>6.2</td>
<td>Results of Teixeira et al.(1975)</td>
<td>104</td>
</tr>
<tr>
<td>6.3</td>
<td>Spatial Domain</td>
<td>107</td>
</tr>
<tr>
<td>6.4</td>
<td>Optimal Control for Sterilization of Pork Puree</td>
<td>109</td>
</tr>
<tr>
<td>6.5</td>
<td>Temperature Distribution at the Center-Plane</td>
<td>111</td>
</tr>
<tr>
<td>6.6</td>
<td>Nutrient Degradation during Cooling Process</td>
<td>113</td>
</tr>
<tr>
<td>6.7</td>
<td>Variation in Maximum Rates of Heating and Cooling</td>
<td>115</td>
</tr>
<tr>
<td>6.8</td>
<td>Nutrient Retention at Different Inlet Hot-Fill Temperatures</td>
<td>117</td>
</tr>
<tr>
<td>6.9</td>
<td>Process Time and Total Sterilization Time at Different Inlet Hot-Fill Temperatures</td>
<td>118</td>
</tr>
<tr>
<td>6.10</td>
<td>Variation in Limits of the Control $u$</td>
<td>119</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.1</td>
<td>Variables and Parameters for System Equations</td>
<td>44</td>
</tr>
<tr>
<td>5.1</td>
<td>System Equations</td>
<td>66</td>
</tr>
<tr>
<td>5.2</td>
<td>Initial, Terminal and Boundary Conditions for System Equations</td>
<td>67</td>
</tr>
<tr>
<td>6.1</td>
<td>Process Data</td>
<td>102</td>
</tr>
<tr>
<td>6.2</td>
<td>Results for Two Dimensional Model</td>
<td>114</td>
</tr>
</tbody>
</table>
## CONTENTS

Abstract 2
Acknowledgements 3
List of Figures 4
List of Tables 6

Chapter 1 Introduction 9

Chapter 2 Principles of Thermal Processing 14
  2.1 Thermobacteriological Considerations 14
  2.2 Retention of Nutrient Quality 18
  2.3 Approaches to Optimization 21

Chapter 3 Optimal Control Theory 25
  3.1 The Minimum Principle 26
  3.2 Bang-Bang Control 32
  3.3 Distributed Minimum Principle 34
  3.4 Remarks 38

Chapter 4 Mathematical Model for Maximizing Nutrient Retention 40
  4.1 Model Formulation 40
  4.2 Optimality Conditions Using the Variational Approach 47
  4.3 Comments on Optimality Conditions 55
  4.4 Constraint on the Retort Temperature 57
Chapter 5 Numerical and Analytical Solutions

5.1 Computational Strategy

5.2 Analytical Solution

5.3 Multiple Switchings: Numerical Solutions

Chapter 6 Two Dimensional Models

6.1 Preliminaries

6.2 Results

6.3 Discussion and Conclusion

6.4 Suggestions for Future Work

Appendix I Nomenclature

References
CHAPTER 1

INTRODUCTION

Sterilization of canned foods is one of the major operations in the food processing industries. The sterilization process involves a suitable heat treatment of canned foods so that the can contents are virtually free of any microorganisms which could degrade the food. It is impractical to achieve complete sterility; the cans are commercially sterilized so that a desired degree of destruction of microorganisms is achieved. One of the major problems in the sterilization is that the thermal processing causes destruction of nutrients and food quality. The deterioration of nutrients and organoleptic properties, and the destruction of microorganisms follow similar kinetics, and hence higher kill rates for microorganisms also involve greater loss in nutrient quality. While it is highly undesirable to incur any degradation of nutrients, it is nevertheless essential to ensure sufficient reduction of microorganisms, usually by a factor of five to ten log cycles, and hence nutrient loss is unavoidable.

Much attention has been paid in recent years to maximizing the nutrient quality in a sterilization process subject to a specified reduction of microorganisms. Sterilization
can be carried out in a batch retort or continuously as in hydrostatic sterilizers. In both cases, the retort temperature and the process time are the two variables which determine the degree of sterilization. Several studies have been made to design a suitable control policy for retort temperatures, with a view to maximizing the nutrient retention. Most of these studies are based on optimizing some functional form for the temperature-time relationship, either analytically or experimentally, and are approximate in nature. However, the recent advances in computer-aided process design, and the development of modern control theory have provided the potential for a new approach to optimization. The basis for this approach is the "minimum principle" which was originally developed for lumped systems by Pontryagin and co-workers in the early sixties (Pontryagin et al., 1962). Subsequent developments have extended the concepts to incorporate distributed parameter systems.

The major advantage associated with the distributed minimum principle is that now the state variables and the control variables can be spatially distributed in any number of dimensions. The control system then involves coupled, nonlinear partial differential equations of first and second order. While the distributed model can provide much more information and can more accurately represent the physical systems than can lumped models, it is also more complex.
Complete analytical solutions are generally not possible. However, with the availability of finite difference and Galerkin finite element methods for discretizing the partial differential equations, numerical solutions can be achieved to a high degree of accuracy.

Application of the distributed minimum principle to food processing problems should provide better procedures for optimization of nutrient quality than those based on less sophisticated approaches.* We make no a priori assumption regarding the functional form (e.g. high temperature short time, ramp, sinusoidal etc.) of the retort temperature versus time relationship. Rather it is the necessary conditions in the minimum principle which determine the optimal temperature profile. Also, the distributed nature of the model allows for a rigorous mathematical formulation of the problem with fewer simplifying assumptions than those embodied in earlier studies. In this thesis, the optimal thermal processing of canned foods is investigated using the distributed maximum principle as a starting point.

In what follows, the salient aspects of food processing operations and the associated kinetics are discussed. In the second chapter, the basics of sterilization processes and the

* Note that the terms 'minimum principle' and 'maximum principle' are used synonymously. Actually, they differ only in sign.
past research work for optimizing nutrient retention are reviewed. The third chapter is primarily concerned with the tools for optimal control. It deals with the continuous time minimum principle, analyzes some of the limitations of the lumped model and establishes the need for a distributed model. A generalized distributed parameters model is also described here.

In the fourth chapter, the emphasis is on the application of the minimum principle to the problem of optimizing nutrient retention. The necessary conditions for optimality are derived here starting from the basic principles of variational calculus. The discussion on numerical methods for solving the above-mentioned system equations, follows in the fifth chapter. We have considered here one dimensional slab geometry for mathematical and computational convenience. Once the nature of the optimal control is known, it is possible to derive analytical solutions for system equations. These are also presented in this chapter. In the sixth and the final chapter, we are mainly concerned with an application of a two dimensional model to sterilization of pork puree, a process for which significant data exists for comparison purposes. Most of the previous research has failed to account rigorously for sterilization during the cooling cycle, necessitating some new definitions. We elaborate two main aspects of our model here, namely the minimum principle leading to optimal nutrient
retention and the distributed nature of the system giving more accurate predictions. Conclusions and some suggestions for future work follow.

Though the present work is primarily concerned with the maximization of the nutrient quality subject to an integral constraint on the final concentration of microorganisms, the results would very well apply to any process which is governed by a diffusion-like equation with the surface concentration or temperature as a control variable. The objective then would be to enhance one kind of reaction product and suppress the other, where both would have similar temperature dependent kinetics with different activation energies.
CHAPTER 2

PRINCIPLES OF THERMAL PROCESSING

It is of prime concern for the food processing engineer to be able to design a flexible heat treatment schedule for sterilizing canned foods. The objective is always to ensure that sufficient destruction of contaminating micro organisms is achieved with a minimum loss in nutrient quality. Since the beginning of the twentieth century, several researchers have reported analytical and experimental work in this area. Bigelow et al. (1920) gave the first general method for calculating thermal process times, and shortly thereafter Ball (1923) defined several parameters of interest in thermal processing and developed mathematical formulas which are still used in the food processing industry. Several others have done pioneering work in this area during the last few decades. There have been numerous textbooks describing the theory and practice of thermal processing, such as those by Leniger and Beverloo (1975), Loncin and Merson (1979), Charm (1978) and Harris and Karmas (1975).

2.1 THERMOBACTERIOLOGICAL CONSIDERATIONS

The micro organisms present in food are destroyed by
thermally induced changes in the original chemical structure of the proteins in the cells. This leads to an inability of the cells to reproduce. The destruction of microorganisms is usually assumed to occur by a first order, irreversible reaction often described by the rate equation

\[ \frac{dN}{dt} = -K N \exp\left(-\frac{E}{RT}\right) \] (2.1)

where \( N \) is the concentration, \( t \) is the time and the term \( K \exp(-E/RT) \) is the temperature dependent reaction rate constant. Chemical reactions occurring in food which lead to degradation of organoleptic quality and destruction of nutrients such as vitamin C and thiamine, follow similar kinetics. The only difference in these cases is that the values of the constants \( K \) and the activation energy \( E \) are much higher for nutrients than for microorganisms. Thompson (1982) and Lund (1975) give a detailed listing of the values of activation energies for different microorganisms and nutrients.

However, it is usual practice in the food engineering literature to denote the temperature dependence of the reaction rate constant by

\[ K = K_{ref} \exp \left[ b(T-T_{ref}) \right] \] (2.2)

so that an equivalent expression for the destruction of microorganisms and nutrients is...
\[ \frac{dN}{dt} = -K_{\text{ref}} N \exp \left[ b(T-T_{\text{ref}}) \right] \]  

(2.3)

Furthermore, microbiologists and food technologists define two more parameters, D and z. D is the time required to reduce the reactant concentration by a factor of 10, at a given temperature, while z is the increase in temperature necessary for reducing the value of D by a factor of 10. From the definition of the reaction rate constant \( K \), one obtains \( D = \frac{2.303}{K} \), so that the kinetic expression becomes

\[ \frac{dN}{dt} = -\left( \frac{2.303}{D_{\text{ref}}} \right) N \exp \left[ \frac{(T-T_{\text{ref}})}{(z/2.303)} \right] \]  

(2.4)

Here the time and the \( D_{\text{ref}} \) values are usually expressed in minutes, and the temperature and the z values in degrees celsius (or Fahrenheit). The reference temperature \( T_{\text{ref}} \) is normally taken as 121.1 degree C (250 F).

Points along the center of a cylindrical can constitute the slowest heating region. The mechanism for heat transfer is assumed to be primarily conduction, though some convective currents may be set up in the case of semi-liquid foods. Leniger and Beverloo (1975) have discussed this aspect qualitatively. The coldest spot may occur below the geometric center of the can if there is no agitation of containers during processing. The process design calculations may be based on the temperature history at these points. Many a time, however, it is desirable to achieve a certain level of average
concentration over the entire can rather than the concentration at a few critical points. Some other factors which can affect the degree of sterilization are: the size and the shape of the cans, movement of cans in the sterilizer, volume of empty space in the sealed can and perhaps the composition of the vapour/gas mixture in this space.

Hayakawa (1978) gives a summary of how the empirical and analytical formulas have been developed in the past. Ball and Olson (1957) have developed tables for the variation in process values with respect to heat transfer coefficients and the difference between the temperature of the center point and the surface of the containers. Stumbo (1953) developed a mathematical model which based the process time on the probability of survival of micro organisms in the whole container, and not just the center of the can. Most recently, Smith and Tung (1982) have reviewed various formula methods and have compared their accuracy. Most of the currently used sterilization methods seem to have a high factor of safety and hence there is considerable scope for saving energy and improving the organoleptic quality.

Lenz and Lund (1977 a,b) have developed a lethality Fourier number method for estimating sterility value at the center of the can. This is an extension of the work done by Lund et al. (1972), which introduces the concept of adiabatic equilibration temperature. The mass average temperature of
the food contents is higher than the centerline temperature during the heating period. Containers are heated until the average temperature inside the can is sufficient to induce the desired kill rates. Then they are held adiabatically until the centerpoint temperature rises and a uniform temperature distribution is achieved. This concept was first applied by Lund for quick blanching operations.

2.2 RETENTION OF NUTRIENT QUALITY

In recent years, research has been directed to the kinetics of nutrient degradation and to the prediction of the nutrient quality of thermally processed foods. Harris and Karmas (1975) and Lund (1977) have discussed the effects of heat processing on nutrients. It may be occasionally possible to improve the taste of the canned food by heat treatment, but that is rarely an objective. In most cases, the nutritional and the organoleptic quality of the food reduces as a result of thermal processing. The kinetics of nutrient degradation are similar to those of the destruction of microorganisms, two parameters of interest being the rate of nutrient destruction $K_r$, at a reference temperature $T_r$, for a first order reaction, and the Arrhenius activation energy $E_a$. These two parameters are equivalent to the $D_r$-value at reference temperature $T_r$, and the $z$ value, discussed in the preceding
It has been observed by Joslyn and Heid (1963) that at higher temperatures, the destruction rate of bacteria accelerates more rapidly than the degradation rate of nutrients, thus favoring high temperature short time (HTST) processes. Ammerman (1957) studied the effects of heat treatment with different processing temperatures and equal microbial lethality values, on selected food constituents, by analyzing the effect on colours, flavours and nutrients such as vitamin C. The two main kinds of microorganisms used as a basis in thermal calculations are C. Botulinum and B. Stearothermophilus. Stumbo (1973) gives the activation energy-values for these microorganisms as 50-80 kcal/mole which are significantly higher than those of nutrients (about 20 kcal/mole). This means that for a given increase in processing temperatures, the rate of microorganism destruction will increase more rapidly than the rate of nutrient degradation. This observation forms the basis for optimizing processes to yield maximum nutrient retention.

The activation energies and the reaction rate constants for different nutrients depend on factors such as pH of the medium, the composition of the food medium and its oxidation reduction potential. There are several quality attributes, e.g. texture, colour and the flavour of food, which exhibit similar degradation due to thermal processing. Their activa-
tion energies vary in the range 10 to 25 kcal/mole (Timbers, 1971). The activation energy for thiamine, a nutrient which has been studied extensively, is not strongly dependent on food medium or composition, and hence the mechanism of thermal degradation of thiamine appears to be the same for different media. However, the reaction rate constant at the reference temperature (or equivalently the D value at the reference temperature) is a function of the medium and the pH. For example, the thiamine destruction at pH 6.6 is about 12 times faster than at pH 3.2. Other nutrients frequently studied include vitamin C and chlorophyll. Their activation energies vary between 10 and 25 kcal/mole.

Mulley et al. (1975 a,b,c) have studied reaction rates for thiamine destruction in various food products and have observed that, in natural foods thiamine is more heat resistant than thiamine in buffered and aqueous solutions, although a first order reaction mechanism seems to hold for both acidic (pH less than 4.5) and non-acidic foods. More recently, Ohlsson (1980 a) and Castillo et al. (1980) have developed quantitative expressions for retention of nutrients.

Thompson (1982) discusses some of the pitfalls of using the simplified approach of first order, irreversible, non-cyclic reaction for predicting nutrient degradation. The degradation rate for a particular nutrient may be a function of the concentration of other nutrients and hence it may
change in a complex manner as various reactant concentrations change during thermal processing. However, in the absence of any detailed and more accurate mathematical modeling for complex reactions, an approximation to first order may be justified. Downes and Hayakawa (1977) and Jen et al. (1971) are among others who have developed mathematical models for predicting nutrient quality. A general treatise on the applicability of reaction kinetics is given by Hill and Grieger-Block (1980).

2.3 APPROACHES TO OPTIMIZATION

There has been growing concern over the nutritional value of canned foods and this has led to active research programmes for optimizing thermal processes. Broadly speaking, optimization refers to a procedure for selecting, among various alternatives, that process which will maximize or minimize some preset quantities, subject to certain constraints. Various approaches have been used for this purpose. For certain heat treatments of canned foods, one may wish to optimize the temperature-time profile, or the size and shape of containers, or the design of equipment. Teixeira et al. (1975) have considered different geometries for containers to yield better heat penetration and thereby to optimize thiamine retention. Flat container geometries or cylindrical cans with
very high height to diameter ratios are favoured. However, in practice the shape and the size of containers is determined by other considerations such as cost of manufacturing and attractiveness of the product.

Processing equipment is designed for providing rotation and agitation of cans during the sterilization process. There are several variations and patented designs in this area. Hydrostatic sterilizers, which can process cans continuously in a pressurized high temperature space, are now used more frequently (Fairbrother, 1982). The optimal retort temperature policy developed in the subsequent chapters, will be applicable to both batch-type processes as well as to continuous sterilizers.

Teixeira et al. (1969) have shown that the maximum retention of nutrients is also dependent on the activation energy or the z value of nutrients. For a low z value (high E) nutrient, a process with low temperature and high time is preferred while for those having high z values, shorter times and higher process temperatures are desirable. Hayakawa and Ball (1971), Teixeira et al. (1975) and Thijssen and Kochen (1980) have considered variable retort temperatures to maximize the nutrient retention. However, one should carefully study the influence of other factors as well as assumptions and simplifications made, if any, before drawing any inference. For example, Teixeira and co-workers have neglec-
ted the sterilization as well as the micro organism destruction during the cooling cycle. This will be discussed in detail in chapter 6. Loncin and Merson (1979) briefly discuss the use of the Euler-Lagrange equation in optimizing nutrient quality.

Saguy and Karel (1979) have used Pontryagin's minimum principle as applied to lumped models, for maximizing the nutrient retention during sterilization of canned foods. This is possibly the only published application of optimal control theory to food processing operations. Figure 2.1 shows the optimal temperature profiles obtained using this technique principle. These temperature policies are shown to improve the thiamine retention by more than two percent as compared with other schemes. Their results are encouraging, although some of their assumptions may not be strictly correct. They have averaged the micro organism and nutrient concentrations over the entire can, and have then used first order kinetics for the averaged quantities. The reason for this averaging seems to be the fact that Pontryagin's minimum principle can be applied only to the lumped models. This is a severe limitation of the model as presented by Saguy and Karel.

There appears to be some scope for the formulation of the optimization problem more rigorously by taking into account the spatially distributed nature of the system, and then applying the minimum principle. This is the prime task of the
thesis research that follows in the subsequent chapters.

![Graph](image)

Optimal retort temperature ($\Theta^*$), mass average temperature ($\Theta_m$), and central point temperature ($\Theta_c$) during the sterilization process (A/2 can).

![Graph](image)

Optimal retort temperature profiles for the sterilization process of pea puree in A303 can and pork puree in A/2½ can.

Fig. 2.1 Results of Saguy and Karel (1979)
Pontryagin's Minimum Principle is one of the major contributions to the development of optimal control theory. It has found a variety of applications in many different types of control problems, especially in electrical and aerospace engineering. In this chapter, we wish to review the basic statement of the minimum principle in one of its commonly used forms, using the state space approach throughout to represent the system. In what follows, we consider a deterministic system with a known relationship between the system states and the input control. The aim is to find the particular control which drives the state $X(t)$ to a desired objective. It is important to assume that all scalar and vector functions are continuous and continuously differentiable to first derivatives (unless otherwise noted). The existence of a solution to the control problem must also be assumed in deriving the necessary conditions. The latter part of the chapter introduces two important aspects related to our subsequent research, namely bang-bang control and the distributed minimum (or maximum) principle.
3.1 THE MINIMUM PRINCIPLE

The minimum principle for lumped models is discussed in this section. Lumped models are characterized by the fact that the state variables are functions of time alone, and that the system dynamics can be described by ordinary differential equations. To begin, assume that the physical process of interest is completely represented by a linear or nonlinear system of coupled or uncoupled differential equations of the form,

\[ \dot{X} = f(X, u, t) \]  (3.1)

where the m-dimensional control vector \(u(t)\) determines the n-dimensional state vector \(X(t)\). A fixed time interval \([t_0, t_f]\) is considered for the process; it will be easy to extend the concepts to the case of variable terminal time. General statements of the initial and final conditions are given by

\[ Q [X(t_0), t_0] = 0 \]  (3.2)

and

\[ R [X(t_f), t_f] = 0 \]  (3.3)

respectively. Obviously, the dimensions of \(Q\) and \(R\) will be less than or equal to \(n\). The following analysis considers the case where initial conditions are specified by equation (3.2).
The objective of the optimal control problem is to search for a control policy \( u(t) \) which will minimize a cost function, \( J \). The exact form of the cost function is a matter of judgement. The optimality conditions, namely the conditions which will lead to a minimization of \( J \), will depend strongly on the nature of \( J \). Hence it is important to be able to quantify the cost or the performance index correctly. Most of the time, the cost function consists of two terms, namely the cost associated with the terminal state, and the accumulated cost over the entire time domain. In mathematical terms, the cost function \( J \) can be written as

\[
J = C [X(t_f), t_f] + \int_{t_0}^{t_f} g[X(t), u(t), t] \, dt \tag{3.4}
\]

Costate variables \( P(t) \), which are similar to the Lagrange multipliers for the case of static optimization, are adjoined to the system equations and incorporated into the cost function such that

\[
J = C [X(t_f), t_f] + \int_{t_0}^{t_f} \{ g[X(t), u(t), t] + P^T(t) \cdot f[X(t), u(t), t] - X(t) \} \, dt \tag{3.5}
\]

A scalar function Hamiltonian is defined as

\[
H[X(t), u(t), P(t), t] = g[X(t), u(t), t] + P^T(t) \cdot f[X(t), u(t), t] \tag{3.6}
\]
Incorporating the Hamiltonian into the cost function, one obtains

\[ J = C[X(t_f), t_f] + \int_{t_0}^{t_f} \{ H[X(t), u(t), P(t), t] - P(t) \dot{X} \} \, dt \]  

(3.7)

A necessary condition for optimality is that the first variation in \( J \) must be zero for independent variations in the state vector \( X \) and the control vector \( u \). This yields the canonical equations for the state and the costate variables (Pontryagin et al., 1962)

\[
\dot{X} = \frac{\partial H}{\partial P} = f(X, u, t) \quad (3.8)
\]

\[
\dot{P} = - \frac{\partial H}{\partial X} \quad (3.9)
\]

and an additional optimality condition,

\[
\frac{\partial H}{\partial u} = 0 \quad (3.10)
\]

The state variables \( X(t) \) are free to end on any terminal manifold (i.e. can have any values at the terminal time). A transversality condition of the form

\[
P(t_f) = \frac{\partial C}{\partial X(t_f)} \quad (3.11)
\]

is obtained for the costate variables. Thus, the terminal condition on the costate variables is determined by the nature of \( C \), the cost for having a particular end point state. The control system equations constitute a two point boundary value
problem. This is an important characteristic of the minimum principle, where the costate variables always evolve backwards in time.

When the terminal time is not fixed, there is an additional equation corresponding to the variation in \( t_f \), which is

\[
H[\dot{X}(t_f), u(t_f), P(t_f), t_f] + \frac{\partial C}{\partial t_f} + (\frac{\partial R}{\partial t_f})^T \lambda = 0 \quad (3.12)
\]

where the end point constraint on the state variables, described by an \( r \) dimensional vector \( R \) [equation (3.3)] is also adjoined to the cost function by an \( r \)-fold Lagrange multiplier \( \lambda \). Note that \( r \) is less than or equal to \( n \). The transversality conditions are then given by

\[
P(t_f) = \frac{\partial C}{\partial X(t_f)} + (\frac{\partial R}{\partial X(t_f)})^T \lambda \quad (3.13)
\]

Thus there are \( r \) additional equations given by (3.3) to account for the \( r \) dimensional unknown vector \( \lambda \).

In cases where the control variable vector is constrained, the necessary condition given by equation (3.10) can hold only if the domain of the optimal control \( u^* \) is completely in the interior of the constraints. Strictly speaking, it is no longer possible to take independent variations in \( u \) and the optimality condition (3.10) must be replaced by the more general form,

\[
H[X^*(t), u^*(t), P^*(t), t] \leq H[X(t), u(t), P(t), t] \quad (3.14)
\]
where the superscript * denotes optimal states. In essence, the necessary condition for optimality is the global minimization of the Hamiltonian $H$ within the constrained domain for $u$.

The variational approach has generally been used in deriving optimality conditions [Berkowitz(1961), Denn(1969)], although it may not always be possible to apply the techniques of variational calculus, for example, in the case of control or state variable inequality constraints. A more rigorous proof of the minimum principle is that given by Athans and Falb (1966). It gives a new interpretation to the costate variables $P$; namely, the costate equation for $P$ describes the motion of a normal of a hyperplane along the optimal state trajectory. The optimal trajectory represents the variation in cost with change in the state space $X$ where time is included as a state variable. According to the principle of optimality, any portion of the optimal trajectory is an optimal trajectory itself. This forms the basis for the proof which considers temporal variations in $t_f^*$ (free terminal time) and spatial perturbations in $u^*$. In effect, it shows that the motion of the hyperplane described by the costate equations does indeed lead to necessary conditions for minimizing the Hamiltonian. The proof of the minimum principle also yields one of the key properties for the state and the costate variables, namely,
\[ \langle \mathbf{x}(t), \mathbf{p}(t) \rangle = \mathbf{x}^T(t).\mathbf{p}(t) = \text{constant} \quad (3.15) \]

i.e. the inner product of the state and the costate variables is invariant with time.

It is not necessary to know the formal proof of the minimum principle to be able to use it. The variational approach is adequate in most cases. Gould (1969) and Ray (1981) have discussed some of the chemical engineering applications of the minimum principle for lumped models (i.e. where the system can be represented by ordinary differential equations). Szepe and Levenspiel (1968) derived an analytical solution for the optimal temperature in the case of catalyst deactivation via first order, irreversible kinetics. Among separation processes, the application to distillation columns has been considered by Robinson (1970) and to cooling tower operations by Jones (1974). Other uses include optimization of polymerization temperature and initial initiator concentration for batchwise radical chain polymerization (Chen and Jeng, 1978), two step catalytic reactions in packed beds (Chang and Reilly, 1976), and the optimal operation of fixed bed reactors in the presence of deactivating immobilized enzyme catalyst (Patwardhan and Sadana, 1982). Yamane et al. (1977) maximized the metabolite yield in fed batch culture using the minimum principle. Murase et al. (1970) have applied Pontryagin's minimum principle in optimizing the
temperature profile for ammonia synthesis in a multitubular reactor heat exchanger system.

3.2 BANG-BANG CONTROL

A very interesting case arises when the Hamiltonian is linear in the control vector \( u \). The necessary condition given by equation (3.10) does not extremize the Hamiltonian with respect to \( u \) since \( \partial H/\partial u \) is now independent of \( u \). It can be immediately seen that if the domain for the control vector \( u \) is infinitely large, then the Hamiltonian can be driven to negative infinity by choosing infinitely large or infinitely small values of \( u \). However, in most physical situations, there will be some bound on \( u \), for example,

\[
\text{u}_{\text{min}} < u < \text{u}_{\text{max}}
\]  

(3.16)

The condition for optimality is given by equation (3.14), which in this case will be

\[
\text{u} = \text{u}_{\text{max}} \text{ for } \partial H / \partial u < 0
\]

and

\[
\text{u} = \text{u}_{\text{min}} \text{ for } \partial H / \partial u > 0
\]  

(3.17)

Thus the Hamiltonian is minimized by operating at the boundaries of the domain of \( u \). The switchings occur at the
singular points where $\partial H/\partial u = 0$. When the term $\partial H/\partial u$ is zero, not just at certain finite points but in a continuous manner for a certain state-space interval, the solution to the problem is said to have a singular arc. Along the singular arc, the contribution of the $u$ term to the Hamiltonian is zero and the Hamiltonian can not be further extremized by any choice of $u$. An additional equation is required to solve for $u$, which is obtained by considering

$$\frac{d}{dt} \left( \frac{\partial H}{\partial u} \right) = 0 \tag{3.18}$$

If the above equation is insufficient, successively higher time derivatives of $(\partial H/\partial u)$ can be set to zero to solve for $u$ in terms of the state and costate variables.

Bellman et al. (1956), Johnson (1965) and Sage (1968) have discussed bang-bang control and the singular solutions arising in optimal control. One of the early applications of the variational approach for deriving necessary conditions for a singular control was given by Desoer (1959). Athans and Canon (1964) have presented a solution in terms of multiple switchings for the problem of minimum fuel consumption in space vehicles, where the control is applied for 'firing', whenever a given amount of consumed fuel would result in the most efficient motion. Farhadpour and Gibilaro (1981) have used bang-bang control for deriving optimal inlet reactant concentrations in the case of an unsteady state operation of a
continuously stirred tank reactor.

Most of the examples of bang bang control, including all those which are mentioned above, are for lumped systems. Many problems are, however, distributed in nature and application of the minimum principle to distributed parameter models is discussed in the next section.

3.3 DISTRIBUTED MINIMUM PRINCIPLE

The emphasis in this section will be on the development of the minimum principle for systems which are distributed in nature. Distributed systems are characterized by partial differential equations in the spatial coordinate vector \( y \) and time \( t \). Butkovskii and Lerner (1960) were among the first to consider the minimum principle as applied to distributed parameter systems.

Consider a distributed system described by

\[
\frac{\partial X(y,t)}{\partial t} = f[\; y, t, u(y,t), \frac{\partial X}{\partial y}, \ldots, \frac{\partial^k X}{\partial y^k}, \ldots]\]

(3.19)

The initial conditions on the state variables are specified at \( t=t_o \), and the boundary conditions on the distributed variables are specified in terms of \( X(t,y) \), \( \frac{\partial X}{\partial y} \), etc. at the boundaries of the spatial domain. The objective is to find an optimal control \( u(y,t) \) which will extremize the cost function
\[ J = \int_{y} c[X(y, \tau_f), \frac{\partial X(y, \tau_f)}{\partial y}, \tau_f] \, dy + \int_{t_{o}}^{t_f} \int_{y} g[X(y, t), \frac{\partial k}{\partial y}, y, u(y, t), t] \, dy \, dt \quad (3.20) \]

As before, the first term in the cost function is the cost associated with the terminal state, while the second term represents the accumulated cost over the entire time and spatial domain. The Hamiltonian \( H \) is defined as before, namely,

\[
H[ X(y, t), \frac{\partial k}{\partial y} Y(y, t), y, t, u(y, t), P(y, t)] = g[X(y, t), \frac{\partial k}{\partial y} Y(y, t), u(y, t), Y, t] + P(y, t) \cdot f[X(y, t), \frac{\partial k}{\partial y} Y(y, t), u(y, t), Y, t]
\]

(3.21)

The system equations are adjoined to the cost function via spatially distributed costate variables \( P(y, t) \) and the cost function can then be expressed in terms of the Hamiltonian. The process time is assumed to be fixed, although extension to a free terminal time is possible. The first variation in \( J \), namely \( \delta J \), is found by taking independent variations in the state, costate and control variables, and one then obtains the following necessary conditions:

\[
\frac{\partial X}{\partial t} = \frac{\partial H}{\partial P} = f
\]

(3.22)
\[ \frac{\partial P}{\partial t} = -\frac{\partial H}{\partial x} - (-1)^k \frac{\partial}{\partial y} \{ \frac{\partial H}{\partial [\frac{\partial x}{\partial y}]} \} \] (3.23)

and

\[ H[x^*(y,t), (\frac{\partial x}{\partial y})^*, y, t, u^*(y,t)] \leq H[x^*(y,t), (\frac{\partial x}{\partial y})^*, y, t, u(y,t)] \] (3.24)

Equation (3.24) is the general form for the global minimization of the Hamiltonian. If the control vector \( u \) is a function of time alone, i.e. \( u(y,t) = u(t) \), then the above condition reduces to

\[ \int_H H[x^*(y,t), (\frac{\partial x}{\partial y})^*, y, t, u^*(t)] dy \leq \int_H H[x^*(y,t), (\frac{\partial x}{\partial y})^*, y, t, u(t)] dy \] (3.25)

Similarly, for \( u(y,t) = u(y) \), this reduces to

\[ \int_{t_0}^{t_f} H[x(y,t), (\frac{\partial x}{\partial y})^*, y, t, u^*(y)] dt \leq \int_{t_0}^{t_f} H[x^*(y,t), (\frac{\partial x}{\partial y})^*, y, t, u(y)] dt \] (3.26)

Note also that when \( u \) is unconstrained, \( \frac{\partial H}{\partial u} \) can be set equal to zero for optimality, either at each point in the case of equation (3.24), or over the entire spatial domain or the time domain in the cases of conditions (3.25) and (3.26) respectively. (Ray and Szekely, 1973)

The distributed components of the state variable vector or its spatial derivatives are generally specified on the
boundaries of \( y \), and the initial state, \( X(y, t_0) \) is assumed to be completely known. The transversality conditions on the costate variables evolve from the variations applied to the state vector \( x \). The end point conditions (\( t = t_f \)) are given as:

\[
P(y, t_f) = C(y, t_f) / X(y, t_f) \tag{3.27}
\]

The transversality conditions on the boundaries of \( y \) are derived for the distributed costate variables as

\[
\partial^{k-1} \left\{ \partial H / \partial \left[ \partial X / \partial y^k \right] \right\} / \partial y^{k-1} = 0 \tag{3.28}
\]

Equation (3.28) holds for all values of \( t \). The necessary conditions for optimality, namely (3.22) to (3.24) and the associated transversality conditions (3.27) and (3.28) are derived using the variational approach. The detailed steps involved in using this approach can be best illustrated for a particular problem at hand. This is shown in the next chapter for the problem of maximizing the nutrient retention of canned foods during sterilization processes.

There have been a few applications of the distributed maximum principle to chemical engineering processes. Hahn et al. (1971) have computed an optimal start up policy for a plug flow reactor wherein a first order, reversible and exothermic reaction takes place. The distributed maximum principle was also applied to the problem of catalyst deactivation by Ogunye and Ray (1971) and by Gruyaert and Crowe (1974).
Nishida et al. (1976) have compared linear and nonlinear lumped models and distributed models, with their application to problems in reaction engineering. They have considered a catalytic series reaction in a tubular, nonisothermal reactor where the catalyst activity is the radially distributed control variable. This appears to be the only example of bang-bang control as applied to a distributed system.

3.4 REMARKS

The necessary conditions derived using the maximum principle, namely equations (3.8) to (3.10) or (3.12) to (3.14) for lumped systems and equations (3.22) to (3.24) for distributed systems, correspond only to local optimality. Actually, the first variation in $J$ yields only the local extremum of the cost function. By taking the second variation $\frac{d^2}{2}J$, it can be shown that this extremum does correspond to a local minimum. However, note that the global minimization of the Hamiltonian does not imply global minimization of the cost function over all values of $u$ (Athans and Falb, 1966). Thus more than one solution may be possible for the given system of state and costate equations and the associated transversality conditions.

Some other observations are also in order. The Hamiltonian $H$ is always minimized as long as the optimal $u^*$ is a
continuous function. This may not occur, for example, at the
switching points in bang bang control, since at these points
the contribution of $u$ to the Hamiltonian is zero irrespective
of the choice of $u$. Thus the Hamiltonian can not be extremized
at the singular points, which is why bang-bang control is
frequently referred to as suboptimal control. Further, it
should be noted that the only difference between the maximum
and the minimum principle is that the Hamiltonian in the
equation (3.14) or (3.24) must be globally maximized rather
than minimized; other necessary conditions remain unchanged.
CHAPTER 4

MATHEMATICAL MODEL FOR MAXIMIZING NUTRIENT RETENTION

The main task in this chapter is to develop a mathematical model for optimal control of the sterilization process. The system equations are cast in dimensionless form, and based on the variational approach, a detailed proof is presented for the derivation of the optimality conditions. It is shown how the problem formulation leads to a bang-bang control policy. The control must be modified when inequality constraints are incorporated into state variables, such as a constraint on the retort temperature. The resulting control policies will be less optimal in such cases than those obtained with fewer or no constraints.

4.1 MODEL FORMULATION

During the sterilization process, heat transfer inside the can occurs primarily by conduction. This holds for solid or semi-solid foods for which the convection currents inside the can are insignificant. Many of the liquid canned foods, on the other hand, have pH less than 4.5, so that sterilization requirements for these foods are not severe. Also, during thermal processing the convective heat transfer coefficients
at the surface of a can are very high, thus enabling one to assume that the temperature at the can surface equals the retort temperature.

Heat conduction inside the can is governed by the equation,
\[
\frac{\partial T(y', t')}{\partial t} = \nabla^2 T(y', t') \tag{4.1}
\]
with appropriate boundary conditions on the time and spatial domains. Thus,

\[
T = T_i \text{ at } t = 0 \tag{4.2}
\]
and, at the planes of symmetry in the spatial domain, we have

\[
\mathbf{q} \cdot \nabla T = 0 \tag{4.3}
\]

where \( \mathbf{q} \) is the unit outward normal vector.

At the external boundaries of the spatial domain,

\[
T = T_R (t') \tag{4.4}
\]

The retort temperature \( T_R(t) \), which is the variable to be controlled, affects the system behaviour only through the boundary condition (4.4). Denn et al. (1966) have discussed solutions to problems where the control may operate only at the boundaries. Here, however, we find it more convenient to apply the minimum principle by transferring the control from
the boundaries to the system equations. For this purpose, a new variable $T_2(\chi',t')$ is defined as follows:

$$T_2(\chi',t') = T(\chi',t') - T_R(t')$$ (4.5)

The heat conduction equation is now modified to read

$$\partial T_2(\chi',t') / \partial t' = \nabla^2 T_2(\chi',t') - dT_R/dt'$$ (4.6)

with the initial condition

$$T_2 = T_1 - T_R(t'=0) \text{ at } t'=0$$ (4.7)

At the planes of symmetry, we have

$$\theta \cdot \nabla T_2 = 0$$ (4.8)

and at the boundaries of $\chi'$,

$$T_2 = 0$$ (4.9)

This formulation transfers the problem inhomogeneity from the boundary conditions to the differential equation itself, and this can have some advantages in the eigenfunction expansion solution of the problem. It is interesting to note that now the rate of change of retort temperature $u'$, or $dT_R/dt'$, should be considered as the control variable rather than the retort temperature itself.

The nutrient and the micro organism concentrations are governed by the equations,
\[ \frac{\partial C_N}{\partial t'} = -K_{o,N}C_N \exp\left(-\frac{E_N}{RT}\right) \quad (4.10) \]

and

\[ \frac{\partial C_M}{\partial t'} = -K_{o,M}C_M \exp\left(-\frac{E_M}{RT}\right) \quad (4.11) \]

respectively. The initial conditions on the nutrient and microorganism concentrations are:

\[ C_N = C_{N0} \text{ at } t'=0 \quad (4.12) \]

and

\[ C_M = C_{M0} \text{ at } t'=0 \quad (4.13) \]

respectively.

Table 4.1 describes the variables and the parameters which appear in the formulation of system equations. In dimensionless form, the system equations can be represented as

\[ \frac{dX_1}{dt} = u \quad (4.14) \]

\[ \frac{\partial X_2}{\partial t} = \nabla^2 X_2 - u \quad (4.15) \]

\[ \frac{\partial X_3}{\partial t} = -a_1 X_3 \exp\left[-\frac{E}{(X_1+X_2+1)}\right] \quad (4.16) \]

\[ \frac{\partial X_4}{\partial t} = -a_2 X_4 \exp\left[-\frac{\beta E}{(X_1+X_2+1)}\right] \quad (4.17) \]

The initial and boundary conditions on the dimensionless variables are given as follows. At \( t=0 \), we have
TABLE 4.1: VARIABLES AND PARAMETERS FOR SYSTEM EQUATIONS.

<table>
<thead>
<tr>
<th>Variables/Parameters</th>
<th>Dimensional</th>
<th>Dimensionless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>$t'$</td>
<td>$t = \alpha t'/ L^2$</td>
</tr>
<tr>
<td>Characteristic Length</td>
<td>$L$</td>
<td>1</td>
</tr>
<tr>
<td>Space</td>
<td>$y'$</td>
<td>$y = y'/ L$</td>
</tr>
<tr>
<td>Temperature</td>
<td>$T(y', t')$</td>
<td>$= X_1+X_2$</td>
</tr>
<tr>
<td>Retort Temperature</td>
<td>$T_R(t')$</td>
<td>$X_1 = (T_R - T_i)/ T_i$</td>
</tr>
<tr>
<td>Initial Temperature</td>
<td>$T_i$</td>
<td>0</td>
</tr>
<tr>
<td>Temperature Difference</td>
<td>$T_2(y', t') = \frac{T - T_R}{T - T_R}$</td>
<td>$X_2 = (T - T_R)/ T_i$</td>
</tr>
<tr>
<td>Nutrient Concentration</td>
<td>$C_N(y', t')$</td>
<td>$X_3 = C_N/ C_{No}$</td>
</tr>
<tr>
<td>Initial Nutrient Concentration</td>
<td>$C_{No}$</td>
<td>1</td>
</tr>
<tr>
<td>Micro organisms Concentration</td>
<td>$C_M(y', t')$</td>
<td>$X_4 = C_M/ C_{Mo}$</td>
</tr>
<tr>
<td>Initial Micro org. Concentration</td>
<td>$C_{Mo}$</td>
<td>1</td>
</tr>
<tr>
<td>Activation Energy</td>
<td>$E_a$</td>
<td>$E = E_a/RT_i$</td>
</tr>
<tr>
<td>Nutrients, Micro org.</td>
<td>$E_N, E_M$</td>
<td>$E = E/(X_1+X_2+1)$</td>
</tr>
<tr>
<td>Exponential Term</td>
<td></td>
<td>$a = K_o L^2/\alpha$</td>
</tr>
<tr>
<td>Reaction Rate Constants</td>
<td>$K_o$</td>
<td></td>
</tr>
</tbody>
</table>

Contd.
<table>
<thead>
<tr>
<th>Variables/Parameters</th>
<th>Dimensional</th>
<th>Dimensionless</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate of Change of Retort Temperature</td>
<td>$u' = \frac{dT_R}{dt'}$</td>
<td>$u = \frac{dX_1}{dt}$</td>
</tr>
<tr>
<td>Thermal Conductivity</td>
<td>$\alpha$</td>
<td></td>
</tr>
<tr>
<td>Ratio of Activation Energies</td>
<td>$\beta$</td>
<td>$= \frac{E_M}{E_N}$</td>
</tr>
</tbody>
</table>
\[ X_1 = X_{10} \quad (4.18a) \]
\[ X_2 = -X_{10} \quad (4.18b) \]
\[ X_3 = 1 \quad (4.18c) \]

and

\[ X_4 = 1 \quad (4.18d) \]

If the initial retort temperature is the same as the inlet hotfill temperature of the cans, then \( X_{10} \) will be zero. Boundary conditions are required only for \( X_2 \). At the planes of symmetry,

\[ \Theta \cdot \nabla X_2 = 0 \quad (4.19) \]

and at the external boundaries of \( y \),

\[ X_2 = 0 \quad (4.20) \]

The system is autonomous, i.e. not explicitly dependent on time \( t \). In shorthand notation, it can be represented as

\[ \frac{\partial x}{\partial t} = f[x, u, \nabla^2 x_2] \quad (4.21) \]

The objective is to maximize the nutrient retention over the entire can during sterilization. This is equivalent to minimizing the cost function \( J \), where
In most sterilization processes, it is essential to achieve a specified reduction in microorganism level. Then the constraint on the average final concentration of microorganisms can be described as

\[(1/V) \int_0^1 X_4 \, dy < C_1 \] \hspace{1cm} (4.24)

where \( V \) is the dimensionless volume of the container. Typical values of \( C_1 \) may be in the range from \( 10^{-5} \) to \( 10^{-10} \).

4.2 Optimality Conditions Using the Variational Approach

The general form of the necessary conditions for optimizing a distributed system was presented in the previous chapter. The aim in this section will be to develop the specific necessary conditions for maximization of nutrient retention by considering the dynamics of the system and using the principles of variational calculus. One-dimensional slab geometry is considered first to keep the mathematics simpler. However, a logical extension to two and three dimensions can be conveniently made. The spatial variable \( y \) has domain \([0,1]\).
The Hamiltonian $H$ is defined as

$$H = \Phi + \bar{p}^T f$$

(4.25)

where $\bar{p}(t)$ is the distributed costate variable vector, $f$ represents the right hand side of the system equation (4.21) and $\Phi$ is the integrand of the cost function defined by equation (4.23). Hence the Hamiltonian can be written as

$$H = a_1 x_3 \exp[-E/(X_1 + X_2 + 1)] + p_1 u + p_2 (\frac{\partial^2 x_2}{\partial y^2} - u) +$$

$$+ p_3 (-a_1 x_3 \exp[-E/(X_1 + X_2 + 1)]) + p_4 (-a_2 x_4 \exp[-\beta E/(X_1 + X_2 + 1)])$$

(4.26)

The state equations (4.14) to (4.17) are adjoined to the cost function via the costate variable vector $\bar{p}$ and the constraint on the final concentration of micro organisms via a Lagrange multiplier $\mathcal{y}$. The modified cost function can then be written as

$$J = \mathcal{y} \left[ \int_0^{t_f} x_4 \, dy - C_1 \right] + \int_{t=0}^{t_f} \left\{ \int_{y=a}^{y=b} \left[ a_1 x_3 \exp[-E/(X_1 + X_2 + 1)] + p_1 (u - dX_1/dt) + p_2 (\frac{\partial^2 x_2}{\partial y^2} - u - \partial x_2/\partial t) + p_3 (-a_1 x_3 \exp[-E/(X_1 + X_2 + 1)] - \partial x_3/\partial t) + p_4 (-a_2 x_4 \exp[-\beta E/(X_1 + X_2 + 1)] - \partial x_4/\partial t) \right] \, dy \, dt \right\}$$

(4.27)
which can be expressed in terms of the Hamiltonian as

$$J = \int \left[ H - \mu \left( \frac{\partial \mathbf{X}}{\partial t} \right) \right] dy dt$$  \hspace{1cm} (4.28)

In the following analysis, the process time $t_f$ is assumed to be a predetermined factor having a fixed value, so that no variations in $t_f$ are considered. Then one can obtain the first variation in $J$ by varying $X, X_{yy}, u$ and $P$. Thus

$$\delta J = \int \left[ \delta X_4 \int_0^t \frac{\partial X}{\partial t} \right] dy + \int_0^t \left[ (\delta X)^T \frac{\partial H}{\partial X} + (\delta P)^T \frac{\partial H}{\partial P} \right. $$

$$+ (\delta u) \frac{\partial H}{\partial u} + (\delta X_{yy})^T \frac{\partial H}{\partial X_{yy}} - P^T \left( \frac{\partial H}{\partial t} \right) dy dt$$  \hspace{1cm} (4.29)

Here $X_{yy}$ has only one non-zero component, namely $X_{2,yy} = \frac{\partial^2 X_2}{\partial y^2}$. Note that

$$P^T \left( \frac{\partial H}{\partial t} \right) = P^T \left( \frac{\partial H}{\partial t} \right) = - (\delta X)^T \frac{\partial P}{\partial t}$$  \hspace{1cm} (4.30)

Also, consider the terms

$$\delta \left[ (\delta X) \frac{\partial \mathbf{H}/\partial X_{2,yy}}{\partial y} \right] / \partial y = [\delta \left( \frac{\partial X}{\partial y} \right) / \partial y].$$

$$[\delta \left( \frac{\partial \mathbf{H}}{\partial X_{2,yy}} / \partial y \right) + (\delta X_2) \frac{\partial^2 \left( \frac{\partial \mathbf{H}}{\partial X_{2,yy}} / \partial y \right)}{\partial y^2}]$$  \hspace{1cm} (4.31)

The first term on the right-hand side of equation (4.31) can be expressed by using another variational relation
Combining equations (4.31) and (4.32) and rearranging, one obtains

\[(Sx_2,yy) \partial_H/ \partial x_{2,yy} = \partial\{(Sx_2,yy)/(\partial x_{2,yy}/\partial y)\} \quad (4.33)\]

The expressions given by equations (4.14) and (4.17) are incorporated into the variation \(SJ\) of equation (4.29) to obtain

\[
\begin{align*}
SJ & = \nu \left( \int_0^t \int_0^t (Sx)^T \partial_H/ \partial x + (Sp)^T \partial_H/ \partial p \right. \\
& + (Su) \partial_H/ \partial u + \partial\{(\partial H/ \partial x_{2,yy}) \partial (Sx_2)/ \partial y\}/\partial y + \\
& \left. (Sx_2)^T \partial_H/ \partial x_{2,yy}/ \partial y^2 - \partial\{(Sx_2) \partial (\partial H/ \partial x_{2,yy})/ \partial y\}/\partial y \right) dy dt
\end{align*}
\quad (4.34)\]

For the first variation in \(J\) to be zero, the coefficients of the independent variations, \((Sp),(Sx)\) and \(Su\) must be zero, which leads to the following necessary conditions:

\[
\partial x/ \partial t = \partial H/ \partial p = \partial
\quad (4.35)\]
\[
\frac{\partial P}{\partial t} = - \frac{\partial P}{\partial X} - \frac{\partial^2 (\frac{\partial P}{\partial X_{yy}})}{\partial Y^2}
\]  
\tag{4.36}

and \[\frac{\partial P}{\partial u} = 0 \tag{4.37}\]

The second term on the right hand side of equation (4.36), namely \(\frac{\partial^2 (\frac{\partial P}{\partial X_{yy}})}{\partial Y^2}\) is non zero only for the variable \(X_2\). Also, strictly speaking, equation (4.37) will be valid only if the variations in \(u\) are allowed to be within the constrained domain for \(u\). Secondly, since \(u\) is a function of time alone, a necessary condition of the type (3.25) will apply. i.e. since the control variable \(u\) does not change over the spatial domain of \(y\), the Hamiltonian \(H\) need be minimized only over the entire domain of \(y\) at any time \(t\), rather than at each point in \(y\). This means that equation (4.37) will modify to

\[
\frac{\partial}{\partial u} \int H dy = \int (\frac{\partial P}{\partial u}) dy = 0
\]  
\tag{4.38}

or more generally,

\[
\int H[x^*(y,t), (\frac{\partial^2 X_2}{\partial Y^2})^*, y, t, u^*(t)] dy = 0
\]  
\tag{4.39}

Next, note that the Hamiltonian as defined by equation (4.26) is linear in the control variable \(u\), so that

\[
\frac{\partial H}{\partial u} = (P_1 - P_2)
\]  
\tag{4.40}
The equation (4.38) cannot yield an optimal value of $u$ since $\frac{\partial H}{\partial u}$ is independent of $u$, and hence a bang-bang control policy is required. The constraint on the rate of rise of retort temperature, $u$, will be of the type

$$u_{\text{min}} < u < u_{\text{max}}$$

(4.41)

where $u_{\text{max}}$ corresponds to the maximum rate of heating and $u_{\text{min}}$ corresponds to the minimum rate of heating (maximum rate of cooling). The desired control policy will then be of the form

$$u = u_{\text{max}} \text{ if } \int_a^b \frac{\partial H}{\partial u} \, dy = \int_a^b (P_1 - P_2) \, dy < 0$$

and

$$u = u_{\text{min}} \text{ if } \int_a^b \frac{\partial H}{\partial u} \, dy = \int_a^b (P_1 - P_2) \, dy > 0$$

(4.42)

Switching(s) in the control will occur at point(s) where the term $I = \int_a^b (P_1 - P_2) \, dy$ is equal to zero. For the existence of a singular arc, the term $I$ must remain zero continuously over a certain time period. This will be discussed in some detail later.

It is readily demonstrated that the Hamiltonian need be minimized only over the entire spatial domain [equation (4.39)] and not at each individual point in the domain of $y$. Consider for example, a spatial discretization of the system equations (4.14) through (4.17). At each time interval a given partial differential equation is converted into a number of coupled ordinary differential equations. If the minimum
principle developed in the preceding chapter is applied to this new, lumped model, an equivalent optimal control can be found. The Hamiltonian at each time interval will be the sum of contributions from each ordinary differential equation. Thus the condition for the global minimization of Hamiltonian will indeed yield equation (4.39) as the number of discretizations goes to infinity.

The transversality conditions are obtained by setting the remaining terms in equation (4.34) for $\tilde{S}J$ to zero. Thus one obtains

$$
\left[ \tilde{S}X_4 \int_{t_0}^{t_f} dy \right]_{t=t_f} + \int_{t_0}^{t_f} \frac{P}{H} \left( \frac{\partial X}{\partial y} \right) \frac{d^2 y}{dt^2} + \left[ \int_{t_0}^{t_f} ( \partial H / \partial X_2, yy ) \cdot ( \frac{\partial X_2}{\partial y} ) \frac{d^2 y}{dt^2} \right]_{t_0}^{t_f} = 0
$$

(4.43)

Terms inside each of the square brackets must be zero, which leads to

$$
P_1(t_f) = 0
$$

(4.44a)

$$
P_2(t_f) = 0
$$

(4.44b)

and

$$
P_3(t_f) = 0
$$

(4.44c)

Also, $P_4(t_f) = \nu$

(4.45)
Since \( x \) is specified at \( t=0 \), \( (\Delta x) \) is uniformly zero at \( t=0 \).

Note that \( \partial X_2 / \partial y = 0 \) at \( y = 0 \) and \( X_2 = 0 \) at \( y = 1 \) so that the transversality conditions on \( P_2 \) become

\[
\partial H / \partial X_{2,yy} = P_2 = 0 \text{ at } y = 1 \tag{4.46}
\]

\[
\partial (\partial H / \partial X_{2,yy}) = \partial P_2 / \partial y = 0 \text{ at } y = 0 \tag{4.47}
\]

Note also that the first necessary condition, equation (4.35), gives back the system state equations. The governing equations for each of the costate variables are:

\[
\partial P_1 / \partial t = -(1-P_3) a_1 X_3 \exp[-E/(X_1+X_2+1)]. E/(X_1+X_2+1)^2 + P_4 a_2 X_4 \exp[-\beta E/(X_1+X_2+1)]. \beta E/(X_1+X_2+1)^2 \tag{4.48}
\]

\[
\partial P_2 / \partial t = -(1-P_3) a_1 X_3 \exp[-E/(X_1+X_2+1)]. E/(X_1+X_2+1)^2 + P_4 a_2 X_4 \exp[-\beta E/(X_1+X_2+1)]. \beta E/(X_1+X_2+1)^2 - \partial^2 P_2 / \partial y^2 \tag{4.49}
\]

which is equivalent to

\[
\partial P_2 / \partial t = \partial P_1 / \partial t - \partial^2 P_2 / \partial y^2 \tag{4.49a}
\]

\[
\partial P_3 / \partial t = -(1-P_3) a_1 \exp[-E/(X_1+X_2+1)] \tag{4.50}
\]

and

\[
\partial P_4 / \partial t = P_4 a_2 \exp[-\beta E/(X_1+X_2+1)] \tag{4.51}
\]
4.3 COMMENTS ON OPTIMALITY CONDITIONS

The complete system of canonical equations is described by equations (4.14) through (4.17) for state variables $X$, equations (4.48) thru (4.51) for costate variables $P$ and the minimization of the Hamiltonian (4.42) for the control $u$. Boundary conditions and initial conditions for the state variables are given as follows.

At $y = 0$, $\partial X_2 / \partial y (0,t) = 0$ \hspace{1cm} (4.19a)

at $y = 1$, $X_2(1,t) = 0$ \hspace{1cm} (4.20)

At $t = 0$,

$X_1(0) = 0$ \hspace{1cm} (4.18a)

$X_2(y,0) = X_{2in}$ \hspace{1cm} (4.18b)

$X_3(y,0) = 1$ \hspace{1cm} (4.18c)

$X_4(y,0) = 1$ \hspace{1cm} (4.18d)

The 'initial' conditions on the costate variables are at time $t = t_f$, and are given by equations (4.44) and (4.45), while the boundary conditions for $P_2$ are given by equations (4.46) and (4.47). Note that the Lagrange multiplier $\varpi$ is an unknown parameter, for which there is an additional constraint equation of the form (4.24) with the equality constraint as
the limiting case, i.e.

\[(1/V) \int_0^t X_4 \, dy = C_1 \]  \hspace{1cm} (4.52)

Some observations are in order. Integrating the equation (4.16) from zero to \( t \), one gets

\[ X_3(y,t) = \exp \left\{ - \int_0^t a_1 \exp \left\{ -E/(X_1+X_2+1) \right\} \, dt \right\} \]  \hspace{1cm} (4.53)

Let \( P_3' = (1-P_3) \), so that the governing equation for \( P_3' \) becomes

\[ \frac{\partial P_3'}{\partial t} = P_3' a_1 \exp \left\{ -E/(X_1+X_2+1) \right\} \]  \hspace{1cm} (4.54)

with

\[ P_3'(t_f) = 1, \]

and integrating the preceding equation from \( t = t_f \) to \( t \), one obtains

\[ P_3'(t) = \exp \left\{ - \int_0^t a_1 \exp \left\{ -E/(X_1+X_2+1) \right\} \, dt \right\} \]  \hspace{1cm} (4.55)

Using the fact that

\[ \int_0^t g \, dt = \int_0^{t_f} g \, dt - \int_0^t g \, dt \]

it is easy to show

\[ P_3'(y,t) = X_3(y,t_f) / X_3(y,t) \]  \hspace{1cm} (4.56)

or \( P_3(y,t) = 1 - X_3(y,t_f)/X_3(y,t) \)  \hspace{1cm} (4.57)
Similarly, it follows that for the state variable $X_4$ and the costate variable $P_4$, we have

$$X_4(y,t) \cdot P_4(y,t) = X_4(y,t_f) \cdot P_4(y,t_f) = f(y) \quad (4.58)$$

Equations (4.57) and (4.58) represent a key property for the costate variables, namely that the inner product of the state and costate variables is time invariant. This is similar to the condition described by equation (3.15) for lumped models. These equations also eliminate the need to integrate two of the costate variables backward in time, or they can be used to provide a check on the accuracy of the numerical integration backward in time for the costate variables.

**4.4 CONSTRAINT ON THE RETORT TEMPERATURE**

Most of the sterilizers in the food industry operate in a certain range between a maximum and a minimum temperature. In many cases, the minimum temperature of the retort is anywhere from 70 to 100 degree C, some times equaling the inlet temperatures for the cans. The maximum temperatures for continuous or batch retorts vary from 125 to 140 degree C, and seldom do they exceed these limits. Clearly, the maximum temperatures at any point inside the can will be well within these limits. Excessively high temperatures must be avoided to prevent localized 'cooking' of the food. It will be essential to
incorporate a constraint on the dimensionless state variable $X_1$, of the form,

$$b_1 \leq X_1 \leq b_2$$  \hspace{1cm} (4.59)

Constraints on the state variables can be treated formally using the penalty-function approach [Bryson and Ho (1975), Sage (1968)]. For the constraint described by equation (4.59), we define a new variable $X_5$, whose dynamics are governed by

$$\frac{dx_5}{dt} = (b_1 - X_1)^2 \cdot \mathcal{H}(b_1 - X_1) + (b_2 - X_1)^2 \cdot \mathcal{H}(X_1 - b_2)$$  \hspace{1cm} (4.60)

where the Heaviside step function $\mathcal{H}$ is defined as

$$\mathcal{H}(f) = \begin{cases} p & \text{for } f > 0 \\ 0 & \text{for } f < 0 \end{cases}$$  \hspace{1cm} (4.61)

It can be immediately seen that $X_5$ is the accumulated penalty for violating the constraint. The quantitative measure of the penalty is given by a positive number $p$ which will have a larger value the more stringent the constraint. If the constraint is to be always satisfied, we should ideally have $X_5(t_f) = 0$. The deviation in $X_5(t_f)$ from zero will represent the penetration into the forbidden domain. Adjoining the system equation (4.60) to the cost function and the Hamilton-
ian of the original problem via a new costate variable $P_5$, one obtains a new Hamiltonian $H'$,

$$H' = H + P_5 \left[ (b_1 - X_1)^2 \mathcal{H}(b_1 - X_1) + (b_2 - X_1)^2 \mathcal{H}(X_1 - b_2) \right] \quad (4.62)$$

The costate equation is given as

$$\frac{\partial P_5}{\partial t} = 0 \quad (4.63)$$

There are no initial or final conditions on $P_5$, while

$$X_5(t=0) = 0 \quad (4.64)$$

and ideally, $X_5(t=t_f) = 0 \quad (4.65)$

From equation (4.63), one can see that $P_5 = \text{constant}$. The value of this constant should be chosen so as to ensure that $X_5$ satisfies equations (4.64) and (4.65).

Figure 4.1 shows a portion of a control policy where one of the switchings occurs at $t=t_s$, for the case of no constraint on the retort temperature. The corresponding control and temperature profiles are shown as dotted lines. Next, consider a case where a constraint of the form (4.59) is imposed on the surface temperature of the can. We can now consider three possibilities, namely $u>0$, $u=0$ or $u<0$ for $t>t_1$ but sufficiently close to $t_1$. For the first case, $u>0$, the retort temperature $X_1$ will be greater than $b_2$ and thus $(X_1-b_2)$ will contribute to the penalty $X_5$, given by equation (4.60). Clearly then $X_5(t_f)$ will be greater than zero. By
Fig. 4.1 Modified Control for Constraint on the Retort Temperature
contraposition, in order for $X_5(t_f)$ to be equal to zero, the constraint must always be satisfied, i.e. $X_t<b_2$. Hence $u$ cannot be greater than zero for $t=t_{1+}$. In fact, this is a trivial example of the use of penalty function analysis. For this case, one could easily argue, without the aid of the penalty function approach, that in order for the constraint to be always satisfied, $u$ must be less than or equal to zero for $t=t_{1+}$. Again as long as the constraint on $X_t$ is satisfied, the Hamiltonian $H'$ defined by equation (4.62) will be the same as the Hamiltonian $H$ for the unconstrained case.

It is implicitly assumed that $u_{min}$ will be less than zero, which must be the case for $u_{min}$ corresponding to the maximum rate of cooling. The control variable $u$ must be chosen so as to minimize the Hamiltonian $H$, defined by equation (4.26) over the entire spatial domain. Note also that the coefficient of $u$ in the expression for $\int H \, dy$, namely the term $I [I = \int (P_1-P_2) \, dy]$, is negative at time $t = t_1$ (point A in figure 4.1). Since the term $I$ is a continuous function, it will continue to be negative for $t>t_1$ but sufficiently close to $t_1$. Hence if $u$ is chosen to be less than zero for $t=t_{1+}$, its contribution to the Hamiltonian will be a positive number and thus the Hamiltonian cannot be minimized by selecting $u$ to be less than zero. This leads to the conclusion that $u$ must be zero after $t=t_1$, which means that the retort temperature must remain constant at $b_2$. This will in turn modify the profile
for the integral term I, so that it may become zero at some time t, not necessarily equal to \( t_s \). The switching to \( u = u_{\text{min}} \) will occur thereafter.

A similar analysis holds for the case when the lower constraint \( X_1 = b_1 \) is approached from above. It should be noted that the contribution of \( u \) to the term \( H \cdot dy \) is uniformly equal to zero when \( X_1 \) is on the boundaries of the constraint. This is however different from the usual case of singular control. In singular control, the contribution of \( u \) to the Hamiltonian is driven to zero because the coefficient of \( u \) is zero along the singular arc. More importantly, in the case of singular arcs the Hamiltonian can not be minimized with respect to the choice of \( u \), as opposed to the above case where \( u = 0 \) is an optimality condition for minimizing the Hamiltonian.

The possibility of singular arcs must not be overlooked either in the above case or in the case where there is no constraint on \( X_1 \). For a singular arc to exist, one must have

\[
\frac{d}{dt} \left[ \int_0^t (P_1 - P_2) \, dy \right] = 0
\]

(4.66)

over a certain range of time. However this is very unlikely to occur since the system state and costate equations are highly nonlinear with several exponential terms and partial derivatives with respect to spatial coordinates. It will be very difficult to foresee a situation wherein the integral term I
is continuously kept zero by simultaneously adjusting different state and costate variables (Tenney, 1982).

The canonical partial differential equations described in this chapter are coupled, nonlinear and distributed, for which a complete analytical solution is not possible. Numerical solutions often involve discretization of the partial differential equations and iteration on the resulting two point boundary value problem. This is dealt with in the next chapter along with a discussion on the results for one dimensional slab geometry.
Optimal control of distributed parameter systems involves the solution of a system of canonical state and costate equations. These are partial differential equations, often coupled and non-linear, with boundary conditions specified at different ends of the time domain for different variables. Clearly, numerical solutions are the only choice for most of these kinds of problems. A procedure for numerical computations is presented in this chapter. This is followed by a closed form solution to the state equations, for the case of bang-bang control with multiple switchings. Numerical solutions for single switching are compared with those for multiple switchings, thereby providing a strong evidence that bang-bang control with single switching is the optimal solution.

5.1 COMPUTATIONAL STRATEGY

It was shown in the preceding chapter that the linearity of the Hamiltonian with respect to the control $u$ leads to bang-bang control where the control now operates only on the boundaries of the admissible domain. We first consider a
solution to the problem described in terms of the one-dimensional slab geometry with no constraints on the state variable $X_1$. It will be possible to extend the ideas further to the case of cylindrical can geometries with surface temperature constraints.

The governing state, costate and control equations, which were derived in the preceding chapter, are presented in collected form in table 5.1. The associated boundary conditions on the state and costate variables are outlined in table 5.2. There are nine variables, namely, $X_1(t)$, $X_2(y,t)$, $X_3(y,t)$, $X_4(y,t)$, $P_1(y,t)$, $P_2(y,t)$, $P_3(y,t)$, $P_4(y,t)$ and $u(t)$; which are to be solved using the nine equations in table 5.1. Note that the state variables $X_1$ to $X_4$ are known at $t=0$, while the costate variables $P_1$ to $P_3$ are specified at $t=t_f$. The costate variable $P_4$ is unspecified at either end. It is equal to an unknown $\varphi$ at $t=t_f$, which evolves from the integral constraint on the final micro organism concentration [equation (4.24)]. The optimal control $u$ is determined by equation (4.42).

Spatial discretization is necessary to convert the given system of partial differential equations to a set of ordinary differential equations. Walsh (1971) and more recently Houstis et al. (1978) have reviewed the finite difference and finite element methods for solving multidimensional partial differential equations. Both finite difference and finite element
TABLE 5.1 SYSTEM EQUATIONS

I. FOR STATE VARIABLES

\[ \frac{dX_1}{dt} = u \] (4.14)
\[ \frac{dX_2}{dt} = \frac{\partial X_2}{\partial y} - u \] (4.15)
\[ \frac{dX_3}{dt} = -a_1 \cdot X_3 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \] (4.16)
\[ \frac{dX_4}{dt} = -a_2 \cdot X_4 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \] (4.17)

II. FOR COSTATE VARIABLES

\[ \frac{dp_1}{dt} = -(1-p_3) \cdot a_1 \cdot X_3 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \frac{E}{(X_1 + X_2 + 1)^2} + \]
\[ p_4 \cdot a_2 \cdot X_4 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \frac{E}{(X_1 + X_2 + 1)^2} \] (4.48)
\[ \frac{dp_2}{dt} = -(1-p_3) \cdot a_1 \cdot X_3 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \frac{E}{(X_1 + X_2 + 1)^2} + \]
\[ p_4 \cdot a_2 \cdot X_4 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \frac{E}{(X_1 + X_2 + 1)^2} - \frac{\partial p_2}{\partial y} \] (4.49)
\[ \frac{dp_3}{dt} = -(1-p_2) \cdot a_1 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \] (4.50)
\[ \frac{dp_4}{dt} = p_4 \cdot a_2 \exp\left[-\frac{E}{(X_1 + X_2 + 1)}\right] \] (4.51)

III. FOR THE CONTROL VARIABLE

\[ u = u_{\text{max}} \quad \text{if} \quad \int \frac{\partial H}{\partial u} \, dy = \int (p_1 - p_2) \, dy < 0 \]
\[ \text{and} \quad u = u_{\text{min}} \quad \text{if} \quad \int \frac{\partial H}{\partial u} \, dy = \int (p_1 - p_2) \, dy > 0 \] (4.42)
TABLE 5.2: INITIAL, TERMINAL AND BOUNDARY CONDITIONS FOR SYSTEM EQUATIONS

(A) Initial Conditions:
At \( t = 0 \),
\[ X_1 = X_{10} \]  
\[ X_2 = -X_{10} \]  \hspace{1cm} (4.18a)
\[ X_3 = 1 \]  \hspace{1cm} (4.18b)
\[ X_4 = 1 \]  \hspace{1cm} (4.18c)

(B) Terminal Conditions:
At \( t = t_f \),
\[ P_1 = 0 \]  \hspace{1cm} (4.44a)
\[ P_2 = 0 \]  \hspace{1cm} (4.44b)
\[ P_3 = 0 \]  \hspace{1cm} (4.44c)
\[ P_4 = \]  \hspace{1cm} (4.45)
\[ \frac{1}{V} \int X_4 \, dy \left[ C_1 \right] \]  \hspace{1cm} (4.24)

(C) Boundary Conditions:
At \( y = 0 \),
\[ \left( \frac{\partial X_2}{\partial y} \right) = 0 \]  \hspace{1cm} (4.19a)
\[ \left( \frac{\partial P_2}{\partial y} \right) = 0 \]  \hspace{1cm} (4.47)
At \( y = 1 \),
\[ X_2 = 0 \]  \hspace{1cm} (4.20)
\[ P_2 = 0 \]  \hspace{1cm} (4.46)
methods were tried in this work. In the finite difference method, the spatial domain was divided into ten equally spaced intervals for the spatially distributed variables $X_2(y,t)$ and $P_2(y,t)$, while for the rest, five equally spaced intervals were used. For the Galerkin finite element approach using bilinear basis functions, the spatial domain was divided into four intervals of width .4, .3, .2 and .1 respectively. Integration in the time domain was done using a fourth order Runge-Kutta-Gill (RKG) method. Details of the RKG algorithm are given by Finlayson (1980).

Brown (1979) gives a detailed discussion on the use of the Galerkin finite element method with different kinds of basis functions. While this code would be highly accurate for two or more spatial dimensions, it offered negligible advantage in terms of accuracy for the case of the one dimensional slab geometry. It requires the solution to a system of algebraic equations at each time step and thus consumes a considerable amount of computer time. The finite difference code, on the other hand, was easier to formulate, took less time for computation and was more accurate with a higher number of meshes than those used for finite elements. Hence in the subsequent analysis, the finite difference code was retained.

A fifth order quadrature formula was developed for averaging the nutrient and micro organism concentrations over
the entire volume. By this formula,

$$\int f \, dy = h \sum_{i=1}^{N} A_i \cdot f(y_i)$$  \hspace{1cm} (5.1)$$

Here $N$ is equal to six for five equally spaced intervals, each of width $h=.2$. The $f(y_i)$ are the values of the function $f$ at each of the nodes. The coefficients $A_i$ were evaluated to be:

$A_1 = A_6 = .329861$

$A_2 = A_5 = .302083$

$A_3 = A_4 = .868056$

This quadrature scheme provides significantly higher accuracy than the lower order formulas such as Simpson's rule or the trapezoidal rule, especially when steep gradients or exponentially decreasing functions are involved.

The general strategy for computation is as follows. The system of canonical partial differential equations is reduced to a set of 53 ordinary differential equations (one for $X_1$; six each for $X_3, X_4, P_1, P_3$ and $P_4$; eleven each for $X_2$ and $P_2$). We begin by assuming a single switching point at $t=t_1$ where the control $u$ changes from $u=u_{\text{max}}$ for $t<t_1$ to $u=u_{\text{min}}$ for $t>t_1$. Equations (4.14) to (4.17) for the state variables $X_1$ through $X_4$ are integrated in the forward direction beginning at $t=0$. Iterations are performed to correct the value of the switching time $t_1$, so that a desired final concentration of the micro
organisms is achieved at \( t = t_f \). This is possible since the state equations are always independent of the costate equations, while the control \( u \) which affects the behaviour of the state equations, is in turn governed by the costate variables, namely equation (4.42). Next, the costate variables \( P_1 \) through \( P_4 \) are integrated backwards in time, beginning with \( t = t_f \). Note that the costate variable \( P_4 \) is unspecified at \( t = t_f \), which is compensated for by the integral constraint on \( X_4 \), namely equation (4.24).

Figure 5.1 shows the plots of the integral term \( I = \int_0^{t_f} (P_1 - P_2) \, dy \) as a function of time, for different values of the unknown \( \nu \), where \( P_4(t_f) = \nu \). (Here and in the subsequent analysis, the time axis is rescaled with respect to the process time \( t_f \).) We find that for large positive values of \( \nu \), the integral is always negative. As we decrease \( \nu \), we observe a singular point (i.e. a point at which the integral \( I \) becomes zero) at \( t = .9 \) for \( \nu = 5 \) and at \( t = .75 \) for \( \nu = 2.5 \). For very small positive values, as well as for any negative values of \( \nu \), the integral is observed to be always positive. One can infer from these results that it is possible to make the term \( I \) equal to zero at a desired value of \( t_1 \) by suitably choosing \( \nu \).

Thus the general scheme for solving the system of equations outlined in table 5.1 would be to integrate the state equations in the forward direction and to iterate for the switching point \( t_1 \) to obtain the desired reduction in micro
Fig. 5.1 Basis for Obtaining a Single Point Switching

- $t_f = 12\text{ min}$
- $E = 25$
- $\beta = 2.5$
- $c_1 = 0.1$
- $a_1 = 3.6 \times 10^{11}$
- $a_2 = 6.956 \times 10^{28}$
organisms at the end of the process. The costate equations are 
then integrated backwards in time to iterate for the value of 
the unknown \( t > t_1 \), 
equal to zero at \( t = t_1 \) and negative for \( t < t_1 \). This will repre-
sent a complete solution to the canonical equations with 
boundary conditions described in table 5.2.

The above approach is a short cut method, which works 
very well partly because of the unique behaviour of the 
costate variables described in section 4.3. A very general 
way of solving a two point boundary value problem is the 
shooting technique. Keller (1972) describes this method for a 
system of two ordinary differential equations (ODE's) where 
two additional ODE's need be integrated in time if one uses a 
Newton-Raphson algorithm to iterate on the unknown initial 
condition for the second variable. This approach could be 
extended to an n-dimensional Newton-Raphson search technique 
in the present case. However, it would result in at least a 
few hundred ordinary differential equations to be solved at 
each iteration, would require a tremendous amount of computer 
time and storage, and again the convergence is not always 
guaranteed. This is a major limitation of the distributed 
parameters model with several states, though much theoretical 
work has been done in this area. The computational strategy 
outlined earlier in this section, appears to be an efficient 
way of solving this problem.
Figure 5.2 shows a plot of the term $I$ for the case of a fixed process time of 24 minutes, with the other parameter values as noted in the diagram. The switching is required after approximately 5 minutes to achieve a $1/10$ reduction in the micro organisms. It is interesting to note that for the case of no constraint on the final concentration of the micro organisms, $P_4(t_f)=0$, which makes $P_4$ uniformly zero everywhere. The results for this case have been shown in figure 5.3, which illustrates that the integral $I$ should always be positive for $t<t_f$, or $u$ should always be minimum for this case. This would indeed make sense, if one were to maximize the nutrient retention with no constraint on the final micro organism concentration.

Once we are convinced that the single switching does represent a possible solution to the given set of equations, we need not be concerned further with the costate equations. This observation is based on the postulate that the condition on the final concentration of micro organisms, uniquely determines the switching point $t_1$, and that there would always be some value of $\nu$ for which the integral term $I$ would have a unique singular point at $t=t_1$. This reduces the computational work to solving only the state equations in the forward direction.

There are two other possibilities. Firstly, it may be possible to have multiple switchings in bang-bang control.
Fig. 5.2 Condition for Optimal Switching

- $E = 25$
- $\beta = 25$
- $c_1 = .1$
- $a_1 = 3.6 \times 10^{11}$
- $a_2 = 6.956 \times 10^{28}$
Fig. 5.3 Condition for Switching for the case of no Constraint on the Final Concentration of Micro Organisms
This is discussed in detail in the subsequent sections. The second possibility is that of a singular arc. Generally in any optimal control problem, it is very difficult to prove mathematically the existence of a singular arc, except in the case of a few, simple, lumped models (Sage, 1968). However in this problem, the presence of singular arcs seems to be very unlikely, since the governing equations contain several non linear, exponentially decreasing terms so that it is very difficult to achieve conditions whereby the integral I would continue to remain zero for a finite time interval (Tenney, 1982).

Figure 5.4 shows the optimal control policy for single point switching with no constraint on the retort temperature. The switching occurs approximately halfway through the process, and the maximum surface temperature corresponds to $X_1 = .12$ (119 degree C). The dimensionless and the dimensional parameters used are:

$a_1 = 3.6 \times 10^{11} , a_2 = 6.956 \times 10^{28}$

$E = 25 , \beta = 2.5 , T_{in} = 350 K$

$u_{max} = (dX_1/dt)_{max} = .25 , u_{min} = (dX_1/dt)_{min} = -.25$

Process time = 30 minutes

$L = 4 \text{ cm} , \alpha = .133 \text{ cm}^2/\text{min}.$
FIG. 5.4 BANG-BANG CONTROL: SINGLE SWITCHING
It is interesting to observe from figure 5.5 how the increase in switching time from $t=0$ to $t=30$ minutes gradually reduces the average micro organism concentration at the end of the process. While a ten log cycle reduction can be achieved by switching about halfway, only an additional log cycle reduction would occur if the heating continues till the end of the process. Clearly, for the fixed process time of 30 minutes in this case the ten log cycle reduction seems best suited. The temperature distribution and the micro organism and nutrient concentrations corresponding to the ten log cycle reduction are shown in figure 5.6.

After the switching occurs, the temperature in the outer portions of the can begins to fall rapidly, while the contents at the center of the can are still at a higher temperature. This leads to the attainment of a desired degree of micro organism destruction in less time than were the can heated uniformly throughout to the retort temperature. Such a control policy is in agreement with the high temperature short time process, followed by a rapid cooling of the products.

In our model, the maximum rates of heating and cooling are considered to be independent parameters, and the optimal solution is found once these limits are known. However, from the nature of bang-bang control, one should really expect an increase in nutrient retention with an increase in the limits of $u$. Figure 5.7 illustrates this behaviour where as much as
Fig. 5.5 Final Concentration of Microorganisms as a Function of Switching Time
Fig. 5.6 Results of Computer Simulation for Bang-Bang Control (Single Switching)
Fig. 5-7 Effect of Variation in the Limits of the Control $u$
50 percent increase in nutrient retention may be observed as one increases the rate of heating and cooling from .225 (3.15 deg C/min) to about .75 (10.5 deg C/min). Correspondingly, the switching time is seen to decrease with a rise in u. Note that the marginal changes in the optimal retention and the switching time decrease rapidly with increasing limits \( u_{\text{max}} \) and \( u_{\text{min}} \).

If we increase the process time, there is scope for further reduction in microorganisms, hence the switching time decreases as shown in figure 5.8, other parameters remaining unchanged. For example, a process time of 35 minutes should be chosen if the energy cost is one of the important considerations.

Different nutrients in the food have different activation energies. For those with lower activation energies, the reaction rate constant is also low, so that the net effect is that of a higher nutrient retention. This is shown in figure 5.9 for the case of ten log cycle reduction of microorganisms.

In many cases, because of the physical constraint, one may observe a lower and an upper bound on the retort temperature, in which case the control policy should be modified as outlined in section 4.4. This leads to the case of dual switching, where the control takes discrete values, \( u_{\text{max}} \), \( u_{\text{min}} \), or zero depending on when the constraint is
Fig. 5.8 Variation in Process Time
Fig. 5.9 Nutrient Degradation at Different Activation Energies
reached. Two typical cases are shown in figures 5.10 and 5.11, where the upper and lower constraints are in effect respectively. Clearly, such a constraint on the retort temperature represents a suboptimal case so that lower nutrient concentrations will be predicted by our kinetic models.

To summarize, we have presented in this section a general computational scheme to solve canonical equations for the case of single point bang-bang control. There are several independent parameters affecting the nutrient retention, such as the process time $t_f$, the maximum rates of change of temperature $u_{\text{max}}$ and $u_{\text{min}}$, activation energies of the nutrients and the micro organisms and finally the constraints on the retort temperature. Effects of each of these parameters has been studied in our computer simulation. Next, we wish to consider analytical solutions to the state equations. Complete closed form solutions are impossible. However, once the nature of the control $u$ is known, the diffusion equation (4.15) can be solved for the case of multiple switching if the switching points are known or can be assumed and iterated upon. This follows in the next section.

5.2 ANALYTICAL SOLUTION

In this section a general solution is developed for the heat conduction equation (4.15), when the control is subject
Fig. 5.10  Optimal Control for Upper Constraint on the Retort Temperature
Fig. 5.11 Optimal Control with Lower Constraint on the Retort Temperature
to multiple switchings. The single point switching or the dual point switching (when there is a constraint on the retort temperature) will be special cases of this solution. We begin with the equation,

$$\partial x_2 / \partial t = v^2 x_2 - u$$  \hspace{1cm} (4.15)

having boundary conditions

$$\Theta \cdot \nabla x_2 = 0 \text{ on the planes of symmetry} \hspace{1cm} (4.19)$$

and

$$x_2 = 0 \text{ on the surface.} \hspace{1cm} (4.20)$$

The corresponding eigenvalue problem will be

$$\nabla^2 v_n = -\lambda_n^2 v_n$$  \hspace{1cm} (5.2)

with boundary conditions

$$\Theta \cdot \nabla v_n = 0 \text{ on the planes of symmetry} \hspace{1cm} (5.3)$$

and

$$v_n = 0 \text{ on the surface.} \hspace{1cm} (5.4)$$

$$v_n(y)$$ is an orthogonal function representing a solution to the differential equation (5.2) with associated boundary conditions (5.3) and (5.4). (See, for example, Myers, 1971.) The general solution to equation (4.15) can be written as
$X_2(y,t) = \sum_{n=0}^{\infty} \langle X_2, V_n \rangle \cdot V_n(y)$ \hspace{1cm} (5.5)

where $\langle a,b \rangle$ denotes an inner product over the entire spatial domain $v$, and is defined by

$\langle a,b \rangle = \int_a \cdot b \, dv$ \hspace{1cm} (5.6)

For the one dimensional slab geometry, we have

$V_n(y) = \sqrt{2} \cos(\lambda_n y)$ \hspace{1cm} (5.7)

and

$\lambda_n = \frac{(2n+1)\pi}{2}$ \hspace{1cm} (5.8)

Multiplying equation (4.15) by $V_n$ and integrating over the volume of the container, one obtains

$\int_v (\partial X_2 / \partial t) V_n \, dv = \int_v \nabla^2 X_2 V_n \, dv - \int_v u \cdot V_n \, dv$ \hspace{1cm} (5.9)

or

$\int_v (\partial X_2 / \partial t) V_n \, dv = \int_v X_2 \cdot \nabla^2 V_n \, dv - \int_v u \cdot V_n \, dv$ \hspace{1cm} (5.10)

which follows on integration by parts twice. With the aid of equation (5.2), equation (5.10) can be reduced to the form

$d\langle X_2, V_n \rangle / dt = -\lambda_n^2 \langle X_2, V_n \rangle - u \langle 1, V_n \rangle$ \hspace{1cm} (5.11)

This equation has a solution
\[<X_2', V_n> = \exp(-\lambda_n^2 t) [<X_2', V_n>_{t=0} - \int_0^t \exp(\lambda_n^2 t') u(t') \, dt']\] (5.12)

where, for the control policies considered here, the integral on the right-hand side is readily evaluated. Thus, if \( t_k \) is the \( k^{th} \) switching time, we have

\[
\int_0^t \exp(\lambda_n^2 t') u(t') \, dt' = \frac{1}{\lambda_n^2} \left\{ \sum_{k=1}^N u_{k-1} [\exp(\lambda_n^2 t_k) - \exp(\lambda_n^2 t_{k-1})] + u_N [\exp(\lambda_n^2 t) - \exp(\lambda_n^2 t_N)] \right\}
\] (5.13)

where \( u_{k-1} \) is the control value in effect during the time interval \((t_{k-1}, t_k)\). Also, \( t_0 = 0 \) and \( N = \max\{k \mid t_k < t\} \). With suitable rearrangement of the terms in the summation in (5.13), the final solution (5.5) can be written as

\[
X_2(x, t) = \sum_{n=0}^{\infty} V_n(x) <1, V_n> [X_{20} \exp(-\lambda_n^2 t) - \frac{1}{\lambda_n^2} \{u_N - u_0 \exp(-\lambda_n^2 t) - \sum_{k=1}^N (u_k - u_{k-1}) \exp(-\lambda_n^2 [t - t_k])\}]
\] (5.14)

The dimensionless retort temperature will be

\[
X_1(t) = X_{10} + \sum_{k=1}^N u_{k-1} \cdot (t_k - t_{k-1}) + u_N \cdot (t - t_N)
\] (5.15)

The nutrient and the micro organism concentrations can be expressed as
\[ X_3(y,t) = \exp \left\{ - \int_0^t a_1 \exp[-E/(X_1+X_2+1)] \, dt \right\} \quad (5.16) \]

and

\[ X_4(y,t) = \exp \left\{ - \int_0^t a_2 \exp[- \beta E/(X_1+X_2+1)] \, dt \right\} \quad (5.17) \]

respectively.

The analytical solution eliminates the need for spatial discretization which is required to solve the system of canonical equations numerically. However, it will now be essential to iterate on the combination of switching points which would give desired final concentration of microorganisms. Substantial numerical integration in time is still required if the analytical solution to the heat equation is used. In the present analysis, we have opted for complete numerical solutions, rather than an analytical plus numerical solution approach. Analytical solutions would be of great help if any instability were to arise in the numerical method due to a very fine discretization in spatial co-ordinates.

5.3 MULTIPLE SWITCHINGS: NUMERICAL SOLUTIONS

It was shown in the earlier part of this chapter that bang-bang control with a single switching point is a possible solution to our problem of optimizing nutrient retention since it satisfies the necessary conditions presented in tables 5.1 and 5.2. It is worth stressing again that there can be more
than one solution to the given set of canonical equations, and each such case would be a suboptimal control. The aim of this section is to explore numerically the possibility of multiple switchings which could constitute such suboptimal cases. The optimal solution would be the one which maximizes the nutrient concentration at the end of the process.

One of the possibilities, not discussed earlier, is that of a zero point switching. In other words, if \( u = u_{\text{max}} \) throughout the process and if there is any upper constraint on the retort temperature, \( u \) will be equal to zero, once this constraint is reached. Such a solution is mathematically possible, if one recalls from figure 5.1 that for very high values of \( v \), the integral \( I \) is uniformly negative. It seems possible to achieve the optimal nutrient concentration at the end of the process with such a control policy; however this is not physically acceptable, since the containers must be cooled down to the room temperature eventually, and hence at least one switching point is necessary.

Next, consider the possibility of two switching points. For this purpose, several combinations of two switching points could be considered. The values of different parameters are the same as those used for the case of single switching point. We begin with the knowledge that for the single switching, \( t_1 = .48 \) as shown in figure 5.4. Obviously, for the case of multiple switching, the first switching point must be less
than this value of $t_1$, if the process begins with $u=u_{\text{max}}$. Figure 5.12 illustrates different combinations of two switching points. As the first switching point $t_1$ (independent variable) moves to the left and away from $t_1=.48$, the second switching point is seen to approach closer to the first switching point, primarily determined by the condition that a desired level of micro organism concentration is achieved at the end of the process. The corresponding retort temperature profiles for the five combinations are also shown. Figure 5.13 shows the results of the computer simulation for one case, where the two switching points are .468 and .664 respectively.

The final nutrient concentration for different combinations of two switching points is shown in figure 5.14. One observes that as the first switching point moves away from the optimal value for single switching, the nutrient concentration continues to decrease. Clearly, the case of single switching ($t_1=.48$), which is the upper most point in figure 5.14, gives the optimal nutrient retention.

Any sterilization process should begin with heating and should end with cooling, hence it is desirable to have $u=u_{\text{max}}$ at the beginning of the process and $u=u_{\text{min}}$ at the end of the process. This will limit us to consideration of only an odd number of switching points. Consider, for example, the case of three switching points, shown in figure 5.15. Here the second
Fig. 5.12 Bang-Bang Control: Two Switching Points
Fig. 5.13 Results for Two Switching Points (Case 1)
Fig. 5.14 Optimal Concentration Using Two Switching Points
Fig. 5.15 Optimal Control: Three Switching Points
switching point $t_2$ is held fixed at .48 and the first switching point $t_1$ is varied from .468 to .42. The third switching point is determined by the integral constraint on the final micro organism concentration. The resulting final nutrient concentrations for different cases are presented in figure 5.16. Again, we find that the nutrient concentration continues to increase as the conditions for optimal single switching are approached. The two plots, obtained with different time step sizes, also illustrate the need for very small time steps for a more accurate integration in time and a correct iteration on switching points. Another interesting behaviour is observed when the second switching time, which was held constant earlier, is changed (figure 5.17). This again leads to a deviation from the optimal conditions for single switching, and hence the resulting nutrient concentrations are lower. Also, the three point switchings seem to give higher nutrient retention than two point switchings in the same range.

Thus, there appears to be strong evidence that single point switching is the optimal control strategy among the various combinations of single, double and triple switchings studied. It can be inferred from the principle of optimality (Athans and Falb, 1966) that, for the case of multiple switchings with more than three switching points, the control trajectory can always be broken down into time intervals in which two or three switchings occur. One could achieve a more
Fig. 5.16 Optimal Concentration with Three Switching Points

Fig. 5.17 Variation in Second Switching Time $t_2$ (3 Switching Points)
optimal nutrient concentration if these two or three switching point controls were replaced by an equivalent single point switching which achieves the same final concentration of microorganisms at the end of these time intervals. Extending this argument further, one can conclude that the single point switching will always be better than any combination of multiple switchings.

Henceforth, we will concentrate only on the single switching control, with or without a constraint on the retort temperature. The next step is to compare the results of the single switching model with the earlier published research work in food sterilization engineering. This will be done in the next chapter, using a cylindrical can geometry.
CHAPTER 6

TWO DIMENSIONAL MODELS

6.1 PRELIMINARIES

The optimal control policy developed in the earlier chapters is applied to the sterilization of canned pork puree in this chapter. For this purpose, we consider a cylindrical can geometry and a five log cycle reduction in the average concentration of micro organisms. Bacillus Stearothermophilus is assumed to be the main food spoilage organism. The Sterilization process is aimed at maximizing the thiamine retention in pork puree. The first order, irreversible kinetics for thermal degradation of nutrients and destruction of micro organisms have been described by

\[ \frac{dN}{dt} = -(2.303 \frac{N}{D_{ref}}) \cdot \exp\left[\frac{(T-T_{ref})}{(z/2.303)}\right] \quad (2.3) \]

as in the work of Teixeira et al. (1975) and Saguy and Karel (1979). The pertinent kinetic data and process parameters listed in table 6.1.

As a starting point, it is worthwhile to estimate the nutrient retention attainable at a certain constant uniform temperature inside the can. This is shown in figure 6.1. It
TABLE 6.1: PROCESS DATA

1. Can dimensions: 307x409 (55/16 inches diameter, 73/16 inches length)
2. Can name: A/2
3. Product: Pork puree
4. Thermal conductivity of pork puree: .09226 cm²/min.
5. Reference temperature: $T_{ref} = 121.1$ degree C (250 F)
6. D value for thiamine at the reference temperature: $D_{ref} = 178.6$ min.
7. z value for thiamine: $z = 25.56$ degree C
8. D value for the micro organism at the reference temperature: $D_{ref} = 4$ min.
9. z value for the micro organism: $z = 10$ degree C
10. Desired reduction of micro organisms: 5 log cycles. i.e. $C_1 = 10^{-5}$

For can specifications, please refer to Lopez, A. (1975).
Fig. 6.1 Maximum Nutrient Retention at Constant Uniform Temperature

t_1 = 89 min
gives an upper bound on the nutrient retention if the maximum retort temperature is known. Obviously, due to conduction resistance, the actual maximum possible thiamine retention will be less than that predicted by the above criterion.

Teixeira et al. (1975) and Saguy and Karel (1979) have investigated optimal temperature profiles for sterilization of pork puree with the process parameters as outlined in table 6.1. They have assumed a constant process time of 89 minutes. Teixeira and co-workers observed that, among different temperature profiles, the ramp gives the maximum nutrient retention. Their results are shown in figure 6.2.

Fig. 6.2 Results of Teixeira et al. (1975)
It is interesting to note that the distributed maximum principle which requires bang-bang control for optimization, leads to the same control policy when there are no active constraints on the retort temperature. The application of the minimum principle to the distributed parameters model explains why one must observe the optimum with the ramp function.

While the results of Teixeira et al. and Saguy and Karel are interesting, the exact values of the maximum nutrient retention should be viewed with caution, since the assumptions made by these researchers concerning sterilization during the cooling process are very different. For example, Teixeira et al. assumed that the destruction of micro organisms is complete when the retort temperature drops below 107 degree C, although the nutrient degradation should continue to occur below this temperature; they did not account for this added degradation in their calculations. Saguy and Karel, on the other hand, have aimed at maximizing thiamine retention at the end of the heating cycle during the sterilization process (which is assumed to be complete when the retort temperature drops below the inlet hot-fill temperature, 71 degree C). They have neglected the nutrient degradation during the cooling cycle, but have accounted for the destruction of micro organisms till the end of cooling.
These discrepancies in the manner in which sterilization and nutrient degradation are considered during the cooling process, lead us to make some more justifying assumptions and to define the terms more precisely, as follows.

1. The destruction of microorganisms and the thiamine degradation is assumed to occur by a first order, irreversible kinetics as described by equation (2.3), which will be valid at least in the temperature range from $T_{\text{max}}$ to $T_{\text{min}}$. $T_{\text{max}}$ can be taken as the maximum retort temperature (130-140 degree C), and $T_{\text{min}}$ as the inlet hot fill temperature which in this case is 71 degree C. It is assumed that there is no further degradation or micro organism destruction below $T_{\text{min}}$.

2. The sterilization process time, or briefly, the process time is defined as the time for thermal sterilization, at the end of which the retort temperature, or more precisely, the temperature at the surface of the can, equals the cooling water temperature, which in our case is taken to be 25 degree C. This change to cooling water temperature may be gradual and continuous in time, or may be a step change.

3. Once the retort temperature equals the cooling water temperature, the additional time required for the temperature at each point inside the can to drop below $T_{\text{min}}$ (71 degree C), will be termed the cooling time. Normally, during cooling, the
retort temperature drops from $T_{\text{min}}$ to the cooling water temperature rather rapidly, so that the cooling time as defined above will always be positive. One could have instead defined the cooling process to begin when the retort temperature drops below $T_{\text{min}}$. However, we will use the first definition. The total process time will comprise the sterilization process time and the cooling time. It is important to note that some sterilization as well as nutrient degradation will be observed towards the end of the process time when the retort temperature drops from $T_{\text{min}}$ to cooling water temperature, and during the cooling time.

Fig. 6.3 Spatial Domain
For the two dimensional can geometry, the conduction equation (4.15) can be written as

\[ \frac{\partial^2 x_2}{\partial t} = \frac{\partial^2 x_2}{\partial y^2} + \left( \frac{1}{y} \right) \left( \frac{\partial x_2}{\partial y} \right) + a^2 \frac{\partial^2 x_2}{\partial z^2} - u \] (6.1)

where \( a \) is the ratio of the diameter to the length of a cylindrical container. The spatial domain is shown in figure 6.3. It is divided into a 10x10 mesh for numerical solution by the finite difference method. The resulting 194 ordinary differential equations are integrated in time using the fourth order Runge-Kutta-Gill algorithm as before.

6.2 RESULTS

From our definition of the process time, it follows that this parameter will not be an independent quantity, but will be determined by the fact that the control is of bang-bang type with \( u \) (the rate of change of temperature) taking the discrete values of \( u_{\text{max}}, u_{\text{min}} \) or zero. Figure 6.4 shows one such optimal control policy where \( u_{\text{max}} \) and \( u_{\text{min}} \) are taken to be equal to 3.87 degree C/min. and -3.87 degree C/min. respectively. The initial retort temperature is 105.4 degree C(222 F) and it rises to a maximum of 129.6 degree C(265 F), whereupon it is held constant. The retort temperature begins to fall after 68 minutes, and is below the inlet hot fill
Fig. 6.4 Optimal Control for Sterilization of Pork Puree
temperature of 71 degree C at the end of 83 minutes. The process time $t_f$ by our definition is 95 minutes. The containers need be held in cooling water for another 32 minutes before the sterilization is complete, i.e. the temperature at each point inside the can is below $T_{\text{min}}$.

Figure 6.5 shows the radial distribution of temperature at the center-plane (i.e. $z=0$) at different times. There is negligible variation in temperature in the vertical direction with steep boundary layers at the upper and lower surfaces. Figure 6.6 shows how the nutrient concentration varies inside the can at the end of the process time $t_f = 95$ minutes and at the end of total sterilization time of 127 minutes. There is a significant drop in the nutrient concentration in the central region of the can during the cooling process. The volume average concentration of nutrient is seen to decrease by about one percent during cooling. The average nutrient retention at the end of total sterilization is 43.6 percent for this case. The volume average temperature at the end of the process time $t_f$ is about 76 degree C. One of the limitations of the lumped models is that the sterilization is assumed to be complete when the mass average temperature drops below $T_{\text{min}}$. However, this is not correct, since the central portions of the can continue to be hotter, causing further degradation and sterilization.
Fig.6.5 Temperature Distribution at the Center-Plane
Nutrient Degradation

Cooling Process

Fig. 6.6 Nutrient Degradation during Cooling Process
It is also not strictly correct to compare the average nutrient retention at the end of a fixed time interval of the heating cycle, since the higher the retort temperature before the completion of the heating cycle, the higher the nutrient retention will be. Refer, for example, to table 6.2, which summarizes the results of the computer simulation for two dimensional models. As shown in the second case, if we choose $u_{\text{max}}$ and $u_{\text{min}}$ to be of magnitude 1.93 degree C/min., the nutrient retention at the end of 89 minutes can be as high as 46.21 percent. (Compare this value with the highest nutrient retention of 45.3 percent as reported by Saguy and Karel, 1979.) However, during the rest of the total sterilization, the average nutrient concentration reduces by 2.6 percent, so that the final nutrient concentration is 43.6 percent.

Figure 6.7 shows the variation in the process time $t_f$ and the total time for sterilization as one increases the magnitude of $u_{\text{max}}$ and $u_{\text{min}}$. With higher rates of heating and cooling, we move closer to step changes in the retort temperature, as a result of which the process time and the total time decrease considerably. On the other hand, the average temperature inside the can at the end of the process time $t_f$, is higher with increased rates of heating and cooling. This leads to a slight increase in cooling time.
TABLE 6.2: RESULTS FOR TWO DIMENSIONAL MODEL

(i) Variation in Heating and Cooling Rates
(ii) Variation in Inlet Hot-fill Temperature

<table>
<thead>
<tr>
<th>Inlet Temp. °C.</th>
<th>Rate of Heating &amp; Cooling °C./min.</th>
<th>&quot;Switching Times, min.</th>
<th>Process Time $t_f$ min.</th>
<th>Cooling Time $t_c$ min.</th>
<th>Total Time $t_f+t_c$ min.</th>
<th>Average Nutrient (%) at $t=t_f$</th>
<th>$t=t_f+t_c$ min.</th>
<th>Average Temp. at $t_f$ °C</th>
<th>Micro org. at $t_f$ X10^5</th>
</tr>
</thead>
<tbody>
<tr>
<td>71</td>
<td>3.87 -3.87</td>
<td>6.23 67.82</td>
<td>94.87</td>
<td>32.22</td>
<td>127.09</td>
<td>44.13</td>
<td>44.79</td>
<td>43.63</td>
<td>75.22 1.26</td>
</tr>
<tr>
<td>71</td>
<td>1.93 -1.93</td>
<td>12.46 64.61</td>
<td>118.72</td>
<td>21.37</td>
<td>140.09</td>
<td>43.67</td>
<td>46.21</td>
<td>43.58</td>
<td>61.31 1.02</td>
</tr>
<tr>
<td>71</td>
<td>5.80 -5.80</td>
<td>4.15 70.04</td>
<td>88.08</td>
<td>36.08</td>
<td>124.16</td>
<td>44.48</td>
<td>44.13</td>
<td>43.57</td>
<td>82.18 3.78</td>
</tr>
<tr>
<td>71</td>
<td>7.73 -7.73</td>
<td>3.12 69.78</td>
<td>83.30</td>
<td>38.18</td>
<td>121.48</td>
<td>44.71</td>
<td>43.86</td>
<td>43.27</td>
<td>87.06 7.17</td>
</tr>
<tr>
<td>88.2</td>
<td>3.87 -3.87</td>
<td>6.23 52.51</td>
<td>79.57</td>
<td>32.57</td>
<td>112.14</td>
<td>46.54</td>
<td>46.11</td>
<td>45.98</td>
<td>75.95 1.68</td>
</tr>
<tr>
<td>105.4</td>
<td>3.87 -3.87</td>
<td>6.23 35.16</td>
<td>62.21</td>
<td>32.58</td>
<td>94.79</td>
<td>53.11</td>
<td>52.55</td>
<td>52.55</td>
<td>75.90 1.40</td>
</tr>
</tbody>
</table>

Average micro organism concentration at the end of total time ($t_f+t_c$), $= 10^{-5}$
Fig. 6.7 Variation in Maximum Rates of Heating and Cooling
It is interesting to note that the inlet hot-fill temperature for the cans is one of the major variables affecting the nutrient retention. Figure 6.8 shows that the average nutrient retention can increase by up to 50 percent as the hot-fill temperature is increased to 100 degree C. Another major advantage in using higher inlet temperature for foods is that the sterilization time is reduced significantly, as shown in figure 6.9. This observation strongly supports the idea that the food contents should be as hot as possible just prior to canning, and the sterilization should be carried out immediately after the canning operation.

The final optimal nutrient concentration varies in a very narrow range from 43.3 percent to 43.6 percent as one enhances the maximum rates of heating and cooling by a factor of four (table 6.2). This is shown in figure 6.10 for simulations done with two different time step sizes. Note that an apparent maximum is observed at $u_{\text{max}} = 4$ degree C/min, which however, flattens out as the time step is reduced by 1/5. The results in table 6.2 correspond to the dimensionless time step .001 (=.089 min). From the nature of the bang-bang control, with higher limits of $u$, a higher nutrient retention should be expected since the Hamiltonian would be maximized by operating at the highest values of the control. This fact is well illustrated for the one dimensional geometry with no constraints on the retort temperature (figure 5.7). The reason
Fig. 6.8 Nutrient Retention at Different Inlet Hot-Fill Temperatures
Fig. 6.9 Process Time and Total Sterilization Time at Different Inlet Hot-Fill Temperatures
Fig.6.10 Variation in Limits of the Control
for not observing this phenomenon in the present case, seems to be associated with the numerical scheme for integration in time domain. With higher rates of heating and cooling, there are large changes in temperatures in a very short time-span, which require very small time steps for accurate integration. This explains the flattening of the nutrient retention curve in figure 6.10. We could expect a marginal increase in nutrient retention, if the minimum step size were further reduced by 1/10 or 1/100. A finer spatial discretization would be another important way to improve the accuracy of the numerical algorithm. However, the excessive amount of computer time and storage required for such integrations, has prevented us from any further reduction in the step size in the time or spatial domain.

Some important inferences can be drawn from the results illustrated in this section. These follow next along with the discussion on other related aspects of the thesis.

6.3 DISCUSSION AND CONCLUSION

The mathematical developments in the preceding chapters have shown that bang-bang control with single-point switching represents the optimal solution to the problem of maximizing the nutrient retention. This is true for a general class of applications where the system dynamics are governed by a
diffusion equation with the control operating on the boundaries, and the cost function and all other system equations being independent of the control. The aim of the thesis was to apply the techniques of optimal control theory to maximize the nutrient retention during thermal processing of canned foods.

The results presented for one dimensional slab geometry will be applicable for the sterilization of cylindrical cans, though numbers may vary. For example, it is easier to show with a one dimensional model that all solutions with multiple switchings represent only suboptimal cases, and that the bang-bang control with a single switching point is the optimal solution. This is true when there are no constraints on the retort temperature, which is rarely found in practice. Again one can show that constant retort temperature with maximum rise and fall of temperature during heating and cooling respectively, is the optimal solution when there are constraints on the retort temperature. This is one of the general practices followed in food process industry.

Several researchers in the past have tried various kinds of functional forms and parametric optimization techniques for maximizing nutrient retention. (See for example, Teixeira et al. (1969 a,b), Thijssen et al. (1978) and Ohlsson (1980 b,c) among others.) Most of this work has been based on trial-and-error optimization, there being no guarantee that a
particular type of functional form or the control policy will be the most optimal. Our approach, on the other hand, does not make a priori assumptions regarding the functional form of the retort temperature; it is rather the necessary conditions derived from the maximum principle which lead us to the optimal control.

Finally, we have modified the model for a two dimensional cylindrical geometry and have applied it to the sterilization of pork puree, on which significant optimization has been done in the past. Our results compare favourably with the highest values of nutrient retention reported by Teixeira et al. (1975), consistent with the fact that the optimal control policy derived by them is a bang-bang control with single switching.

The work of Saguy and Karel (1979) served as a starting point for this research. They have shown how Pontryagin's minimum principle can be applied to the problem of maximizing nutrient retention. However, many of their assumptions were less rigorous than those made here, and this work extends their analysis in a logical manner.

One of the major contributions of our model lies in its distributed nature, so that one can compute temperatures and nutrient and micro organism concentrations accurately. More importantly, there are no errors resulting from the lumping of these quantities. This has also helped us to rigorously define
the process time and the cooling time, and to account correctly for the sterilization during the cooling process, which was not done in the work of others. With the example of the sterilization of pork puree, we have illustrated that our optimal control policy will lead to maximum nutrient retention under given conditions.

Last but not the least, it is interesting to note that inlet hot-fill temperature of the canned foods is one of the major variables, affecting the nutrient retention. Hence every effort should be made to keep this temperature as high as possible prior to sterilization.

6.4 SUGGESTIONS FOR FUTURE WORK

The last point mentioned in the preceding section gives us some clues to the direction of future research. There is a significant scope for further improving the nutrient retention and at the same time, reducing the process time, by modifying the on-line processes prior to sterilization, to ensure high inlet hot-fill temperatures. This research is likely to foster more in an industrial environment.

In this thesis, we have concentrated on conduction heated foods. However, there are several kinds of semi-solid foods as well, wherein the convective currents set up inside the can will enhance the heat transfer. Normally, the industrial
sterilizers are designed to provide end-to-end, parallel axis or other type of rotation and agitation of cans, so that convection can occur inside the can. While much experimental work has been done in this area, no one has yet developed a sound theoretical model for predicting the convective heat transfer coefficients.

A substantial portion of this thesis has been devoted to the mathematical development of the distributed minimum principle. One can find a vast scope for theoretical as well as applied research in this field. However, from our point of view, the techniques of optimal control are a means to an end, and not an end in themselves.
### APPENDIX I

### NOMENCLATURE

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_i$</td>
<td>coefficients in the quadrature formula (i=1 to 6)</td>
</tr>
<tr>
<td>$C$</td>
<td>cost associated with the terminal state</td>
</tr>
<tr>
<td>$C_M$</td>
<td>micro organism concentration, no./cm$^3$</td>
</tr>
<tr>
<td>$C_{M0}$</td>
<td>initial concentration of micro organisms, no./cm$^3$</td>
</tr>
<tr>
<td>$C_N$</td>
<td>nutrient concentration, g/cm$^3$</td>
</tr>
<tr>
<td>$C_{N0}$</td>
<td>initial concentration of the nutrient, g/cm$^3$</td>
</tr>
<tr>
<td>$C_f$</td>
<td>final concentration of micro organisms</td>
</tr>
<tr>
<td>$D$</td>
<td>can diameter, cm</td>
</tr>
<tr>
<td>$D_r, D_{ref}$</td>
<td>time required to reduce the reactant concentration by a factor of 10, min.</td>
</tr>
<tr>
<td>$E$</td>
<td>dimensionless activation energy</td>
</tr>
<tr>
<td>$E_M$</td>
<td>activation energy for destruction of micro organisms, kcal/mol</td>
</tr>
<tr>
<td>$E_N$</td>
<td>activation energy for nutrient degradation, kcal/mol</td>
</tr>
<tr>
<td>$E_a$</td>
<td>activation energy, kcal/mol</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>H</td>
<td>Hamiltonian</td>
</tr>
<tr>
<td>I</td>
<td>integral term ((P_1 - P_2) \ dy)</td>
</tr>
<tr>
<td>J</td>
<td>cost function</td>
</tr>
<tr>
<td>K, K_0</td>
<td>reaction rate constant, (\text{min}^{-1})</td>
</tr>
<tr>
<td>L</td>
<td>characteristic length: slab thickness for one dimensional geometry, can height for two dimensional geometry</td>
</tr>
<tr>
<td>L</td>
<td>Lagrange multiplier vector</td>
</tr>
<tr>
<td>N</td>
<td>concentration, (\text{mol or no./cm}^3)</td>
</tr>
<tr>
<td>P</td>
<td>costate variables vector</td>
</tr>
<tr>
<td>P_i</td>
<td>costate variable associated with the state variable (X_i) ((i = 1) to (5))</td>
</tr>
<tr>
<td>Q</td>
<td>vector expressing initial conditions</td>
</tr>
<tr>
<td>R</td>
<td>gas constant, (\text{kcal/mol.K})</td>
</tr>
<tr>
<td>R</td>
<td>vector expressing terminal conditions</td>
</tr>
<tr>
<td>T</td>
<td>temperature, (\text{(^{\circ}C) or K})</td>
</tr>
<tr>
<td>T_i, T_{in}</td>
<td>initial temperature of the can, (\text{(^{\circ}C) or K})</td>
</tr>
<tr>
<td>T_{max}</td>
<td>maximum temperature for the retort, (\text{(^{\circ}C})</td>
</tr>
<tr>
<td>T_{min}</td>
<td>minimum temperature below which no further nutrient degradation or micro organisms destruction occurs. (\text{(^{\circ}C})</td>
</tr>
<tr>
<td>T_R</td>
<td>retort temperature, (\text{(^{\circ}C) or K})</td>
</tr>
<tr>
<td>T_{r, ref}</td>
<td>reference temperature, K</td>
</tr>
<tr>
<td>T_2</td>
<td>temperature differential ((T - T_R)), (\text{(^{\circ}C) or K})</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>V</td>
<td>volume of the can</td>
</tr>
<tr>
<td>( V_n )</td>
<td>characteristic functions in the eigenfunction expansion</td>
</tr>
<tr>
<td>X</td>
<td>state variables vector</td>
</tr>
<tr>
<td>( X_1 )</td>
<td>dimensionless retort temperature</td>
</tr>
<tr>
<td>( X_2 )</td>
<td>dimensionless equivalent of ( T_2 )</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>dimensionless nutrient concentration</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>dimensionless micro organisms concentration</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>penalty for violating the constraint</td>
</tr>
<tr>
<td>( X_{10} )</td>
<td>initial value of ( X_1 )</td>
</tr>
<tr>
<td>( X_{20} )</td>
<td>initial value of ( X_2 )</td>
</tr>
<tr>
<td>a</td>
<td>diameter to length ratio for the can</td>
</tr>
<tr>
<td>( a_1 )</td>
<td>reaction rate constant for nutrient degradation</td>
</tr>
<tr>
<td>( a_2 )</td>
<td>reaction rate constant for destruction of micro organisms</td>
</tr>
<tr>
<td>b</td>
<td>term in the kinetic expression (2.2), equivalent of activation energy</td>
</tr>
<tr>
<td>( b_1 )</td>
<td>lower constraint on the retort temperature</td>
</tr>
<tr>
<td>( b_2 )</td>
<td>upper constraint on the retort temperature</td>
</tr>
<tr>
<td>f</td>
<td>right-hand side of the system equations in a vector form</td>
</tr>
<tr>
<td>( f(x) )</td>
<td>function of ( x )</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>( f(x_i) )</td>
<td>value of the function at ( x_i )</td>
</tr>
<tr>
<td>( g )</td>
<td>time variable cost term</td>
</tr>
<tr>
<td>( h )</td>
<td>element width in spatial discretization</td>
</tr>
<tr>
<td>( t )</td>
<td>time</td>
</tr>
<tr>
<td>( t' )</td>
<td>time, min.</td>
</tr>
<tr>
<td>( t_c )</td>
<td>cooling time</td>
</tr>
<tr>
<td>( t_f )</td>
<td>process time</td>
</tr>
<tr>
<td>( t_i )</td>
<td>( i^{th} ) switching time</td>
</tr>
<tr>
<td>( t_0 )</td>
<td>initial time</td>
</tr>
<tr>
<td>( u )</td>
<td>control variable</td>
</tr>
<tr>
<td>( u )</td>
<td>control variable vector</td>
</tr>
<tr>
<td>( u' )</td>
<td>control variable, ( ^\circ \text{C}/\text{min.} )</td>
</tr>
<tr>
<td>( u_{\text{max}}, u_{\text{max}} )</td>
<td>maximum value of the control</td>
</tr>
<tr>
<td>( u_{\text{min}}, u_{\text{min}} )</td>
<td>minimum value of the control</td>
</tr>
<tr>
<td>( y )</td>
<td>spatial co-ordinate</td>
</tr>
<tr>
<td>( \mathbf{y} )</td>
<td>spatial co-ordinate vector</td>
</tr>
<tr>
<td>( y' )</td>
<td>dimensional spatial co-ordinate, cm</td>
</tr>
<tr>
<td>( z )</td>
<td>increase in temperature necessary to reduce the D value by a factor of 10</td>
</tr>
<tr>
<td>( z )</td>
<td>spatial co-ordinate along the vertical axis of the can</td>
</tr>
<tr>
<td>Symbol</td>
<td>Definition</td>
</tr>
<tr>
<td>--------</td>
<td>------------</td>
</tr>
<tr>
<td>( \alpha )</td>
<td>thermal conductivity, ( \text{cm}^2/\text{min.} )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>ratio of the activation energies ( E_M/E_N )</td>
</tr>
<tr>
<td>( \theta )</td>
<td>unit outward normal vector</td>
</tr>
<tr>
<td>( \psi )</td>
<td>value of ( P_4 ) at ( t=t_f )</td>
</tr>
<tr>
<td>( \phi )</td>
<td>integrand of the cost function</td>
</tr>
<tr>
<td>( \mathcal{H} )</td>
<td>heaviside step function</td>
</tr>
<tr>
<td>( \lambda_n )</td>
<td>eigenvalue in the solution by eigenfunction expansion</td>
</tr>
</tbody>
</table>
REFERENCES


See also,


130


see also,


Gruyaert, F. and Crowe, C.M. 1974. On the optimization of
distributed parameter systems with boundary control: A counter example for the maximum principle. AIChE J. 20:1124.


133


Stumbo, C.R. 1953. New procedures for evaluating thermal
processes for foods in cylindrical containers. Food Technol. 7:309.


