Book

Universal Optimization and its Application

Alexander Bolonkin

(Excerpts from book)

Abstract

The book consists of three parts. The first part describes new method of optimization that has the advantages at greater generality and flexibility as well as the ability to solve complex problems which other methods cannot solve.

This method, called the "Method of Deformation of Functional (Extreme)", solves for a total minimum and finds a solution set near the optimum. Solutions found by this method can be exact or approximate. Most other methods solve only for a unique local minimum. The ability to create a set of solutions rather than a unique solution has important practical ramifications in many designs, economic and scientific problems because a unique solution usually is difficult to realize in practice.

This method has the additional virtue of a simple proof, one that is useful for studying other methods of optimization, since most other methods can be delivered from the Method of Deformation.

The mathematical methods used in the book allow calculating special slipping and breaking optimal curves, which are often encountered in problems of optimal control.

The author also describes the solution of boundary problems in optimization theory.

The mathematical theory is illustrated by several examples. The book is replete with exercises and can be used as a text-book for graduate courses. In fact the author has lectured on this theory using this book for graduate and post-graduate students in Moscow Technical University.

The second part of the book is devoted to applications of this method to technical problems in aviation, space, aeronautics, control, automation, structural design, economic, games, theory of counter strategy and etc. Some of the aviation, aeronautic, and control problems are examined: minimization of energy, exact control, fuel consumption, heating of re-entry space ship in the atmosphere of planets, the problems of a range of aircraft, rockets, dirigibles, and etc.

Some of the economic problems are considered, for example, the problems of a highest productivity, the problem of integer programming and the problem of linear programming.

Many economic problems may be solved by the application of the Method to the Problems of non-cooperative games.

The third part of the book contains solutions of complex problems: optimal thrust angle for different flight regimes, optimal trajectories of aircraft, aerospace vehicles, and space ships, design of optimal regulator, linear problems of optimal control.

This book is intended for designers, engineers, researchers, as well as specialists working on problems of optimal control, planning, or the choosing of optimal strategy.

For engineers the book provides methods of computation of the optimal construction and control mechanisms, and optimal flight trajectories.

In addition, the book will be useful to students of mathematics, general engineering, and economic.

TABLE OF CONTENTS

Chapter III. MaxiMin method

Chapter IV. Numerical implementation of some algorithms α-functional and maximin. Other numerical methods.

Part Two. APPLICATION OF METHODS *α*-, *β*- functionals and Maximin for technical problems

Part 1

Mathematical Base of the Optimization Methods

Abstract

A new method of optimization by means of a redefinition of the function over a wider set and a deformation of the function on the initial and additional sets is proposed.

The method (a) reduces the initial complex problem of optimization to series of simplified problems, (b) finds the subsets containing the point of global minimum and finds the subsets containing better solutions that the given one, and (c) obtains a lower estimation of the global minimum.

Introduction

The classical approaches this problem is following:

Problem A. *Find a minimum of the given function.*

Together with problem A the following problems are considered:

Problem B. *Find a smaller subset contains the all points of the global minimum.*

Problem C. *Find a subset of better solutions where the function is less that given value.*

Problem D. *Find a lower estimation of function.*

These non-classical approach B,C, and D require innovative methods, different from the well-known methods.

The author offers a new mathematical methods for the solution of these problems.

The new methods have turned out to be much more general, so that while solving one of the above problems, another may be solved in passing, which may help in the solution of the former. Thus, if a satisfactory lower estimate found, it can be compared with various engineering solutions and give rise to one very close to the optimum.

This method is applied to many mathematical problems of optimization. For example, functions of several variable, constrained optimization, linear and nonlinear programming, multivariable nonlinear problems described by regular differential equations and equations in partial derivatives, etc.

One can easy get from the given method to many well-known methods of optimization, for example, Lagrangian multiplier method, the penalty function method, the classical variational method, Pontragin's principle of maximum, dynamic programming and others.

At present, the most of researchers in optimization fields are using the traditional optimization problem – find a minimum of the given functional (Problem A). They look a single, local minimum. An engineer, however, is usually interested in a subset of quasi-optimal solutions. He must make sure that the optimum does not exceed a given value (Problem C). Also, a good estimation from below will indicate how far a given solution is from the optimum solution (Problem D). An addition an engineer usually has other considerations that cannot be introduced into a mathematical model or can lead to impractical complications. Approach C provides him with some choice.

Problem D is also of particular interest. If an estimate from bottom closes to the exact infinum of the function is found, the optimization can frequently be reduced to finding a quasi-optimal solution by trial and error.

Solution of the Problem B can significantly simplify the solution of any of the above problems, since it narrows the set containing optimal solution.

These non-classical Problems B.C. and D require innovative methods, different from the well-known method of variational calculus, maximum principle and dynamic programming. This new method is general, so that while solving one of the above problems, another may be solved in passing, which may help in the solution of the former. Thus, if a satisfactory estimate from below has been found, it can be compared with various engineering solutions and give rise to one very close to the optimum.

Our reasoning in this book is not complex. But we are using symbolic of set Theory, which many engineers forget. That way we are given these information in Appendix A of the book.

In Book we are using the double numbering of formulae, theorems and drawings. The first figure in nubbering formule or theorem notes the number of paragraph, the second figure is number formula or theorem in this paragraph. The first figure of drawings means the number of chapter, the second is the number of drawing.

Chapter 1

Methods of β and γ functions

§1. Methods of β −functions

1. Statement of the Problem. Main theorems. Algorithm 1.

1⁰. Statement of the Task. Assume that the state of the system is described by element *x*. A series of these elements form the set *X*={*x*}. The numerical function *I(x)* (functional) is defined and bounded by its lower estimate over *X*. The relationships and limitations imposed on the system yield a subset $X^* \subseteq X$.

Traditionally the problem of optimization has been set as follows:

A. Find a point of the minimum of the function *I(x)* over the set *X**.
We shall also consider the following problems:

B. Find a smaller subset $M \subset X^*$ that contains point *x** of global (absolute) minimum, $x^* \in M$.
C. Find a subset $N \subset X^*$ on which *I(x)≤ c*, where *c ≥ I(x).*
D. Find the lower estimates of *I(x)* over *X**.

We will name the point (element) *x* the solution if *x* is result any presses, procedure, calculation or reasoning. It not means that *x* is point of optimum. We will tell the point x_1 is better solution than the point x_2, if $I(x_1) < I(x_2)$ and the point of the same solution, if $I(x_1)=I(x_2)$.

For simplicity we assume that the point of global minimum *x** exists in *X**, but this is not impotent limitation. The most results can be obtained without this assumption.

Let us introduce a set *Y={y}* and define a bounded numerical function (functional) *β(x,y)* over $X{\times}Y$. We shall call it ***β-functional***.

Then we set

$$J(x, y) = I(x) + \beta(x, y).$$

Call our initial problem of finding *x** and $I(x^*) = \inf\ I(x) = m, \quad x \in X^*$ **Problem 1**

and the problem of finding $\bar{x}$ and

$$\bar{J}(\bar{x}(y), y) = \inf[I(x) + \beta(x, y)], \quad x \in X$$ **Problem 2**

We assume that $\bar{x}(y)$ exist over $X{\times}Y$.

We deformed arbitrarily our functional *I(x)* by adding *β(x,y*). Moreover we widened the domain of the deformed functional and arbitrarily defined it on the set *Y*. we should do so in such a way that problem 2 will be easier to solve.

It might seem that this makes no sense because we must find the points of minimum of our initial functional *I(x)*, i.e., solve Problem 1. But it appears that from the solution of the simpler Problem 2 we can obtain information about Problem 1. We can use freedom in choice of the functional *β(x,y)* and the set *Y* for such a deformation of functional *J(x,y)* and the set *Y* that we solve the initial Problem 1, but in an easier way.

2^0. The Fundamental Theorem. The following main theorem establishes the relationship between Problem 1 and 2, as well as between Problems A, B, and C (The Principle 1 of Optimum).

Theorem 1.1. *Distinguishing between the sets containing: (1) The global minimum points, (2) only better solutions than the one given, (3) only worse solutions than one given.*

Assume $X^* \equiv X$, $\bar{x}(y)$ *are the points of global minimum in Problem 2. Then:*

(1) The points of global minimum in Problem 1 are contained in the set

$$M = \{x : \beta(x,y) \geq \beta(\bar{x}(y), y),\ \ y \in Y\};$$

(2) The set

$$N = \{x : J + I \leq \bar{J} + \bar{I}, \quad y \in Y\}$$

contains the same or better solutions (that is over N, we have $I(x) \leq I(\bar{x})$ *);*

(3) The set

$$P = \{x : \beta(x,y) \leq \beta(\bar{x}(y), y),\ \ y \in Y\};$$

contains the same or worse solutions (that is over P $I(x) \geq I(\bar{x})$ *).*

<u>Proof.</u> 3. By subtracting the inequality $\beta(x,y) \leq \beta(\bar{x}(y), y)$ *from* $I(x) + \beta(x,y) \geq I(\bar{x}(y)) + \beta(\bar{x}(y), y)$ *we get* $I(x) \geq I(\bar{x})$ over *P*. Point 3 of the theorem is proved.

1. Point 1 of Theorem 1 is obvious because *X=M+P* and $I(x) \geq I(\bar{x})$ over *P*, we have $x^* \in M$. Point 1 of the theorem is proved.
2. By subtracting the inequality $J \geq \bar{J}$ *from* $J + I \leq \bar{J} + \bar{I}$ *we get* $I(x) \leq I(\bar{x})$ over *N*. Point 2 of the theorem is proved.
   Theorem 1 is proved.

If in sets *N* and *P* we write the strong inequality $\beta > \bar{\beta}$, then the set N will contain only better solutions and the set P will contain worse solutions that $I(\bar{x})$.

Theorem 1.1 is correct when *X*≠X*, but *M,N,P* contain elements from *X**.

Let us focus our attention on the fact that after solving the simpler Problem 2, we distinguished in our set *X* three subset: *M*, which contains a point of global minimum, subset *P*, containing the same or worse solutions, and subset *N*, which contains the same or better solutions.

Consequences:

1. Element $\bar{x}$ is the point of global minimum of the functional over the set *P⊆X*.

2. $\bar{x}$ is the element which gives the maximum of the functional *I(x)* over the set *N⊂X*.

3. If *X*⊆P*, then $\bar{x}$ is the point if global minimum Problem 1 over set *X**. In this case we have *M={x}.*

4. If *β=β(x), x∈X, then*

$$M=\{x:\beta(x)\geq\beta(\bar{x})\},\qquad P=\{x:\beta(x)\leq\beta(\bar{x})\},\qquad N=\{x:J+I\geq\bar{J}+\bar{I}\quad .$$

Theorem 1 is correct when *X*≠X*, but *M,N,P* contain element from *X**.

5. Let *X*≠X*. If *X*∩M=∅,* then $I(\bar{x})$ is the lower estimation *I(x)* over the set *X** (because in this case we have *X*⊆P*).

6. Let *X*≠X*. If *X*⊂N*, then $I(\bar{x})$ is the top estimation *I(x)* ≤ $I(\bar{x})$ over the set *X**.

If $\bar{x}$ ∈*X**, the sets *M,N,P* will always contain at least one element from the set *X**. This element is $\bar{x}$.

Remarks:

1. *N⊆M*. The proof: Let us denote $\overset{o}{P}$ =*P*-{$\bar{x}$}. Then $\overset{o}{P}$ ∩ *N*=∅, because over $\overset{o}{P}$ we have *I(x)>I(*$\bar{x}$*)* and over *N* we have *I(x) ≤ I(*$\bar{x}$*).* But *N⊂X* and *M=X*-$\overset{o}{P}$. Hence *N⊆M*.
2. Assume the definitions of the sets *N*, *P* (see Theorem 1) contain strong inequalities. Then the set *N* will contain on; *y* better solutions and the set *P* – only worse solutions, compared to $\bar{x}$.
3. We can use the dependence of the sets *M,N,P* from *y* in order to change the "dimensions" of these sets.
4. *β* - functions exist and their number is infinite.
   The last statement is obvious because we can define *β*-functionals over the set *X×Y* in any possible way.

The theorem 1 gives the Algorithm 1 (a *β*-functional method for finding the subsets that contains the points of global minimum or better solutions).

Algorithm 1. *Define* $\beta_i(x,y)$ *so that Problem 2 becomes easier to solve, and find sets* M_i *and* N_i*. Then* $M=\cap M_i$ *(that is not empty) is the set that contains the points of global minimum and* $N=\cap N_i$ *(if that is not empty) is subset contains min*$\{I(\bar{x}_i)\}$ *or better solutions.*

Note: The getting *M* is more "narrow" (contains less points *x*) subset then initial *M.* That means the finding *x** is easier. The decreasing of *M* is especially important in a "method of dynamic programming" because it is decreasing the number of computation.

Theorem 1.2. (The lower estimate) *Let us assume that β(x,y) is a defined and bounded functional over X×Y then the lower estimate over X is*

$$I(x)\geq[I(\bar{x}(y))+\beta(\bar{x}(y),y)-\sup_X\beta(x,y)]\quad\text{for }\ \forall y\in Y\,. \tag{1.1}$$

Proof: By adding the inequalities

$$I(x)+\beta(x,y)\geq I(\bar{x})+\beta(\bar{x}(y),y)\quad \mathit{and}\quad -\beta(x,y)\geq-\sup_X\beta(x,y)$$

over *X*, we get the estimate (1.2).

Remarks:

1. For case *β = β(x)* the estimate (1.1) is

$$I(x) \geq \inf_{X} J(x) - \sup_{X} \beta(x) \text{ ,} \tag{1.1'}$$

2. When $X \neq X^*$ the estimate (1.1) is correct over *X**, because $X^* \subseteq X$. In this case we can use the better estimates:

$$I(x) \geq \inf_{X^*} J(x) - \sup_{X} \beta(x), \quad I(x) \geq \inf_{X} J(x) - \sup_{X^*} \beta(x), \quad I(x) \geq \inf_{X^*} J(x) - \sup_{X^*} \beta(x), \tag{1.1''}$$

When we found the set *M* for β_i the following estimate may be used

$$I(x) \geq \inf_{X^*} J(x) - \sup_{M} \beta(x), \tag{1.1'''}$$

The proof of (1.1'), (1.1''), (1.1''') is same the proof of theorem 1.2.

3. Dependence of the estimate (1.1) from *y* may be used for its improving

$$I(x) \geq \sup_{y}[\inf_{x^*} J(x) - \sup_{x} \beta(x)], \tag{1.1^{IV}}$$

When we use the estimates (1.1') - (1.1^{IV}) we decide the problem $\widehat{\beta} = \sup_{X} \beta$. It may be used for

Theorem 1.3. Assume *X=X**, $\overline{x}$ is point of a global minimum in the problem $\widehat{\beta} = \sup_{X} \beta$,

Then:

1) The points of global minimum in Problem 1 are contained in the set

$$M(y) = \{x : I + \beta \leq \widehat{I} + \widehat{\beta}, \quad y \in Y\}$$

Contains the same or better solutions.

2) The set

$$N(y) = \{x : \beta - I \geq \widehat{\beta} - \widehat{I}, \quad y \in Y\}$$

3) The set

$$P(y) = \{x : I + \beta \geq \widehat{I} + \widehat{\beta}, \quad y \in Y\}$$

Contains the same or worse solutions.

Here is $\widehat{I} = I(\widehat{x})$.

Proof of Theorem 1.3.

1, 3. By subtracting the inequality $\beta \le \widehat{\beta}$ from $I + \beta \ge \widehat{I} + \widehat{\beta}$ we get $I \ge \widehat{I}$ over set *P*.
The statement 1, 2 follow from this.

2. By subtracting the inequality $\beta \ge \widehat{\beta}$ from $I - \beta \ge \widehat{I} - \widehat{\beta}$ and multiply this result by -1,
we get $I \le \widehat{I}$ over N. The theorem 1,3 is proofed.

Remark:

For proof of the theorems 1.1-1.3 the existence of x, $\overline{x}$, $\widehat{x}$ is not important, but corresponding *inf* and *sup* must be existed.

Example 1.1.

Find minimum of functional

$$I = -e^{-x^4}\cos x^2 - \frac{0.1}{x^2 - 0.2x + 1}, \quad -\infty < x < \infty, \tag{1.2}$$

Solution. Take

$$\beta(x) = \frac{0.1}{x^2 - 0.2x + 1}.$$

Then

$$J = I + \beta = e^{-x^4}\cos x^2.$$

The minimum of this *J* is obvious: $\overline{x} = 0$.

From theorem 1.1 we got the point of the global minimum is in set

$$M=\{x: \beta(x) \ge \beta(0)\}$$

or

$$\frac{0,1}{x^2 - 0.2x + 1} \ge 0.1 \ ,$$

The solution of this inequality is

$$0 \le x \le 0.2 \ .$$

It's not difficult to find the point of global minimum in this small interval by any known method.

We get the lower estimate (theorem 1.2)

$$J(0)-\sup_x \beta = -1-0.101 = -1.101.$$

Value *I(0)* = -1.100 . We see *I(x)* for $x = 0$ is very close to global minimum.

Example 1.2

Find minimum

$$I = -\frac{0.1}{x^2-2x+10} + \cos 4\pi x - 4\cos 2\pi x, \quad -\infty < x < \infty \tag{1.3}$$

Solution: We take

$$\beta(x) = -\cos 4\pi x + 4\cos 2\pi x \ .$$

Then

$$J = I + \beta = \frac{0.1}{x^2-2x+10}, \quad \bar{x} = 1.$$

This solution is global minimum of Problem1 over set

$$P = \{x: \beta(x) = \beta(1)\}$$

or

$$-\cos 4\pi x + 4\cos 2\pi x \le 3.$$

We transform this inequality in

$$-8\sin^4 \pi x \le 0 \ .$$

We see $P = \{x: |x| < \infty\}$. Therefore $P = X^*$. That means (see Consequence 1) $\bar{x} = 1$ is point (and alone) of global minimum of the functional (1.3).

Example 1.3 .

More full, we are demonstrating the new method on following simple functional.

Find the absolute minimum of the functional

$$I = 2x^4 + x^2 - 2x + 1 \ \text{ on the set } X^* = \{x: |x| < \infty\} . \tag{1.4}$$

It is a simple example, which can be solved using well-known methods. For example, take the first derivative, make it equal to zero. Solve an algebraic 3-d order equation (it may not be a simple task) and then analyze the points so found with respect to maximum and minimum.

We shall try to solve this example by the above method as it follows from algorithm 1.

Let us introduce a series $\beta_i(x)$. As follows from Theorem 1.1 we have the sets M_i:

1) Take $\beta_1=2x$. Then
$J = I + \beta_1 = 2x^4 + x^2 + 1, \quad \bar{x} = 0, \quad \text{from} \quad \beta \geq \bar{\beta} \quad \text{we} \quad \text{have} \quad M_1 = \{x : x \geq 0\}$.
As we see the domain which contain a global minimum have become less in two times.

2) Take $\beta_2= -x^2+2x$. Then
$J = I + \beta_2 = 2x^4 + 1, \quad \bar{x} = 0, \quad \text{from} \quad \beta \geq \bar{\beta} \quad \text{we} \quad \text{have} \quad M_1 = \{x : 0 \leq x \leq 2\}$.

Our interval contained a global minimum is only $0 \leq x \leq 2$.

For given β_2 we can use an estimation of the functional which follows from Theorem 1.2.

$$I(x) \geq J(\bar{x}) - \sup_X \beta_2(x) = 1 - \sup_X(-x^2 + 2x) = 1 - 1 = 0,$$

where the point of supreme of β is $\hat{x} = 1$.

From theorem 1.3 we have the additional set M:

$$M_3 = \{x : J(x) \leq J(\hat{x})\} \quad or \quad M_3 = \{x : |x| \leq 1\}.$$

As we see the set $M = M_2 \cap M_3 = \{x : 0 \leq x \leq 1\}$, The global minimum of this problem is in the interval $0 \leq x \leq 1$.

3) Take $\beta_3 = 2x^2 + 2x - 0.5$. Then $J = I + \beta_3 = 2x^4 - x^2 - 0.5$. From inf J we have $\bar{x}_{1,2} = \mp 0.5$.
4) Find for point x_1 set M:

$$\bar{x}_1 = -0.5, \quad M_4 = \{x : -0.5 \leq x \leq 1.5\},$$

$$\bar{x}_2 = 0.5, \quad M_5 = \{x : 0.5 \leq x \leq 0.5\}.$$

The estimation gives $I(x) \geq 3/8 - 0 = 3/8$.

We see that the diameter of the set $M=\cap M_i$ decreases until reduces in the point $\bar{x} = 0.5$. Therefore this point is one of the absolute minimum of the Problem 1 and $I(0.5) = 3/8$.

5^0. The geometric illustration of Theorem 1.1 is given in fig, 1.1 for single variable. The curves $I(x)$, $J(x)$, $\beta(x)$, $I(x)+0.5\,\beta(x)$ and point $\bar{x}$ are drawn. There are the sets M, N, P. P is set x, where

$\beta(x) \leq \beta(\bar{x})$, M is set $X\backslash P$ and N is set x, where $J(x) + 0.5\beta(x) \leq J(\bar{x}) + 0.5\beta(\bar{x})$.

We can see that $N \subset M$.

In fig.1.2 we see sets M, N, P for the case when $I(x_1,x_2)$ is function of two variables x_1 and x_2.

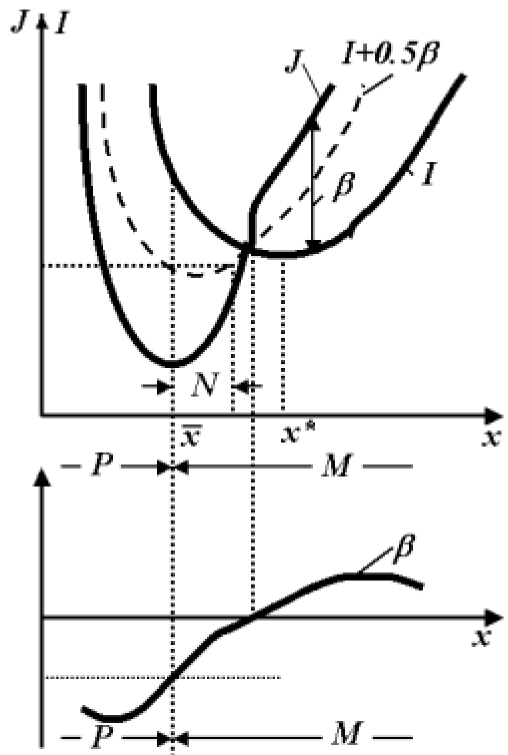


Fig. 1.

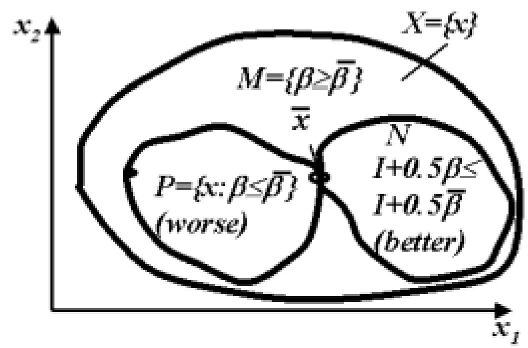


Fig. 2.

Fig.1.1. Geometric illustration of Theorem 1.1 for case of single variable.

Fig.1.2. Sets *M, N, P* for case of two variable.

2. About Convergence of Algorithm 1.

Consider condition of convergence $\inf_{x\in X} J(x), \quad x\in X \quad to \quad \inf_{x\in X^8} J(x), \quad x\in X^* \quad and \quad \bar{x} \quad to \quad x^*$ for Algorithm 1. when we have the succession $\beta_i(x)$, *I* = 1,2,… This succession gives the succession of the sets M_i, N_i and values of functionals $J(\bar{x}_i)$.

The succession $\{\inf J(\bar{x}_i)\}$

For $i \to \infty$ is monotonous decreasing and bounded of bottom, that's way it has a limit. If this limit equals one of lower estimates, that $J(\bar{x}) = I(x^*)$.

Let us to consider now convergence of diameter *d(M), d(N)* of sets $M=\cap M_i$, $N=\cap N_i$ for $i\to\infty$.

This convergence is also monotonous decreasing and bounded of bottom: $d \geq 0$. Therefore it has a limit.

We have got the following simply criterian of convergence

Theorem 1.4. *Assume, the point of the absolute minimum of functional I(x) over set X=X* is single.*

If $d(M)\to 0$, than $x=\lim M(i)=x^*$, $i\to\infty$.

In this case the set contained of point of global minimum $M=\cap M_i$ decrease in point. Therefore this point is the point of the absolute minimum of Problem 1.

Let us take succession of function $W_s(x)$, s =1,2,... . Take $\beta_i(x)$ as

$$\beta_i = \sum_{s=1}^{i} c_s W_s(x) \tag{1.5}$$

where c_s is constants.

We will take these constants c_s from condition

$$\Delta_i = \min_c [I(\bar{x}_i) - \inf_X J_i(x) + \sup_X \beta_i(x)].$$

The value ΔI is difference functional from its lower estimate. Other words value Δ show how much value $I(\bar{x}_i)$ differs from optimum. We name this number Δ-estimate (*delta-estimate*). It is obvious that succession $\{\Delta_i\}$ is monotonous decreasing because every next sum (1.5) contains previous sum. It is also limited of bottom ($\Delta_i \ge 0$). Therefore the succession $\{\Delta_i\}$ converge.

From destination Δ_i we get the following

Theorem 1.5. *If* $\Delta_i \to 0$ *Than* $\inf_X J(x) \to \inf_{X^*} I(x)$.

Theorem 1.6. *Assume X=X*,* $\beta_i = c_i\beta(x)$, *I(x), β(x) is continuous and β(x) is limited on X.*

Then, if $c_i \to 0$ *we have* $J(x) \to m = \inf I(x)$ *over X*.*

Statement of Theorem 1.6 follows from continuous *J(x).*

This theorem may be useful for finding of the local minimum of *I(x)* by way of methods of successive approximations. Assume c_1=1 and problem inf *J(x)* can decided simply. Because functional *J(x)* is continuous, we can wait, that small change of c give small changing (moving) $\bar{x}$.

Therefore $\bar{x}$ is good the initial approximation for $c_2 < c_1$. It is known, that a good initial approximation is very important for speed of convergence. We come to *x** by decreasing *c* to 0.

These criterions of convergence may be used for solutions Problem *A, B, C*, D (see §1,A).

3. Modification of the Theorem 1.1

Over we have considered the case, when we are looking for the additional function $\beta(x,y)$ such us the problem 2 became simpler for solution.

But sometimes it's more comfatable to take such function $J(x,y)$ that the problem $\inf_X J(x,y)$ became easy for solution.

In this case Theorem 1.1. better to write as following

Theorem 1.1'. *Assume* $X^{\bullet} \equiv X$, $\bar{x}(y)$ *is the point of global minimum in Problem 2 .*

$$\bar{J} = \inf_X J(x.y)$$

Then

1) *The points of global minimum in Problem 1 are contained in the set*

$$M(y) = \{x:\ J - I \geq \bar{J} - \bar{I},\quad y \in Y\}.$$

2) *The set*

$$N(y) = \{x:\ J + I \leq \bar{J} + \bar{I},\quad y \in Y\}$$

Contains the better or same solutions.

3) *The set*

$$P(y) = \{x:\ J - I \leq \bar{J} - \bar{I},\quad y \in Y\}$$

Contains worse or same solutions.

This Theorem is correct if $J = kJ_1$, where k = const>0.

4. Method of big steps in set of better solutions. Algorithm2.

From the Theorem 1.1 we can get the following

Algorithm 2 (Method of big steps in set of better solutions)

Take any point x_1 from X^* and such function $J_1(x)$ that point x_1 is its minimum. Find the set N_1 of better solutions. Take from this set a point x_2 and such function $J_2(x)$ that x_2 is its minimum. Find the set N_2 and so on.

It is obvious that $N_1 \supseteq N_2 \supseteq N_3 \supseteq ...$. Let us suppose that result of this process is following - set N_i become point x_N .

Theorem 1.7. *Assume X* is open set, I(x), J_i(x) are continuously and differential (of Freshe) on X*.*

Then point x_N is a stationary point of the function I(x) over X.*

Proof in Appendix 4° of Chapter 1.

Theorem 1.8. *If in point x_N we have*

$$\beta(x_N) - I(x_N) = \sup_{X^*}[\beta(x) - I(x)],$$

Then x_N is point of global minimum of Problem 1.

Proof is in Appendix 5° of Chapter 1.

If conditions of Theorem 1.8 is executed only in small sphere around point x_N then x_N is point of local minimum of Problem 1.

The example for illustration of this method (for tests of constrained minimum) will be given in § 4 (remark 4.3).

We can get the direction in the set *N*, if we calcule a gradient of function in *N.*

The advantegies this method with comparison of gradient method is big steps. When you are in set *N*, you have not a danger of to get worthier solution than given one. This can substentionaly decrease amount of calculation.

5. Method of β-function for Problems with constrains

A) Assume *I*(*x*) is function by its lower estimate over set *X*. The subset $X^* \neq \varnothing$ is separated from *X* by functions

$$F_i(x) = 0 \quad i = 1,2,..., k, \quad \Phi_j(x) \le 0, \quad j = 1,2,..., q, \tag{1.6}$$

where *x* - is *n*-dimentional vector of numerical values.

Take β-function as following (we have a sum for lower index *i,j*)

$$\beta(x, y) = \lambda_i(x, y)F_i(x) + \omega_j(x, y)\Phi_j(x),$$

where $\lambda_i(x,y)$, $\Phi_j(x, y)$ are functions of *x,y*, $y \in Y$, $\omega_j(x, y) \ge 0$.

Write *J*-function

$$J(x,y) = I(x) + \lambda_i(x,y)F_i(x) + \omega_j(x,y)\Phi_j(x) \,. \qquad (1.8)$$

Theorem 1.9. *Assume exist x* ∈X*, y is fixed.*

In other $\bar{x}$ *to be a point of global minimum of function I(x) over X* necessary and enough to exist of function β(x,y) such as*

$$1)\ J(\bar{x},y) = \inf_{x\in X} J(x,y), \quad 2)\ \bar{x}\in X^*, \quad 3)\ \omega_j(x,y)\geq 0 \quad over \quad X, \quad 4)\ \beta(\bar{x},y)=0\,, \qquad (1.9)$$

The **proof** in Appendix 6° of Chapter 1.

Theorem 1.10. (The lower estimation)

Assume y is fixed, $\bar{x}$ *is point of minimum (1.8) for conditions* $\omega_j(x,y) \geq 0$.

Then $J(\bar{x},y)$ *is lower estimation of function I(x) on X*.*

Proof: On set *X** we have $\lambda_i F_i \equiv 0, \quad \omega_j \Phi_j \leq 0$ (that is $\beta(\bar{x},y) \leq 0$). Since over *X** we have $J(\bar{x},y) \leq I(x)$. Theorem is proved.

Likely a common case for β - function we can get the sets

$$M = \{x : \beta \geq \bar{\beta}\}, \quad N = \{x : J + I \leq \dot{J} + \bar{I}\}, \quad P = \{x : \beta \leq \bar{\beta}\}$$

and in this case.

Freedom in choice of *y* we can use for improvement of estimation and decrease sizes of sets *M, N*. Remark only that $\bar{x} = \bar{x}(y)$ and for every *y* corresponding $\bar{x}$ you must find inf *J*(*x,y*), *x*∈*X*.

Remark:

We can take β-function (1.7) in form

$$\beta(x) = \frac{1}{2} a \sum_{i=1}^{k} F_i^2(x) + \sum_{j=1}^{q} a^{\Phi_j(x)} \,.$$

It is possible to show for some conditions: [*I*(*x*), $\Phi_j(x)$, $F_i(x)$ are continuous, *x* is compact set, *x** is close set and don't contain separated points; *x** ∈*X** and exist], when $a \to \infty$, we have $\bar{J} \to m, \quad \bar{x} = x^*$.

B) Assume $F_i(x) = 0$ in (1.6) absent, i.e. the Problem is

$$I(x) = \min, \quad \Phi_j(x) \leq 0, \quad j = 1,2,\ldots,q$$

For solution of this problem we can use following **algorithm**:

1. Take any functions ω (*x,y*) (it's may be less zero) and find the point $\bar{x}(y)$ of global minimum (one may be implicit form $\xi(\bar{x}, y) = 0$) of general numerical function

$$J = I(x) + \sum \omega_j(x, y)\Phi_j(x) \quad on \quad X \ . \tag{1.12}$$

2. Solve equations

$$\xi(\bar{x}, y) = 0, \quad \omega_j(\bar{x}, y)\Phi_j(\bar{x}) = 0, \quad j = 1,2,..., q \tag{1.13}$$

3. Select from these solutions such which satisfy inequalities

$$\omega_j(\bar{x}, \bar{y}) \geq 0, \quad j = 1,2,..., q \ . \tag{1.14}$$

These are points of global minimum of Problem (1.11) because all request the theorem 1.4 is satisfy.

We can solve (1.13) by different ways. For example, find $\bar{x}$ from equation $\xi(\bar{x}, y) = 0$ and substitute in the last equations (1.13)

$$\omega_j(\bar{x}(y), y)\Phi_j(\bar{x}(y)) = 0, \quad j = 1,2,..., q \tag{1.15}$$

Find *y* from this system of equations. Select from these solutions such which satisfy inequalities

$$\omega_j(\bar{x}(y), y) \geq 0, \quad j = 1,2,..., q \ , \tag{1.16}$$

or we can find *y* from $\xi(\bar{x}, y) = 0$ and substitute in the last equations (1.13) and find $\bar{x}$.

6. Application the method of β - functions to linear programming.

The Problem of Linear Programming is

$$I = \sum_{i=1}^{n} c_i x_i = \min, \quad \sum_{j=1}^{n} a_{kj} x_j - b_k \leq 0, \quad k = 1,2,..., m \tag{1.17}$$

Here c_i, a_{kj}, b_k are constant.

Take $\omega_j = y_i$. Then equation (1.13) are

$$y_k (\sum_{j=1}^{n} a_{kj} x_j - b_k) = 0, \quad k = 1,2,..., m \tag{1.18}$$

$$c_i + \sum_{j=1}^{m} a_{ij} y_j = 0, \quad i = 1,2,..., n \tag{1.19}$$

Selective from (1.18) l equations $(l \le n,\ l \le m,\ l = \max)$ and l variables x_j such that determinant $|a_{kj}| \ne 0$. Find $\tilde{x}_j$ from these l linear equations (1.18) (corresponded $y_k \ne 0$).

If this solution don't satisfy inequalities (1.17), we take l other equations and repeat this procedure (process) while we find $\tilde{x}_j$ which saticfy (1.17). If these equations absent, we take l -1 equations (1.18) and repeat process, than l - 2 equations and so on, while we get l = 0.

If solution, which satisfy (1.17), absent that inequality (1.17) is conflicting (incompatible) and cannot be solved.

Assume that by using this procedure we find the solution $\tilde{x}_j$, that satisfy (1.17). Take in (1.19) all y_j, which don't belong the taken questions (1.18), equal zero and find y from equation (1.19). If all $\tilde{y}_j \ge 0$ then $\tilde{x}_j$ is point of minimum of problem (1.17). If part of $\tilde{y}_j < 0$, then we change corresponded equations (1.18) by other and repeat this process while get all $\bar{y}_j \ge 0$.

We can suppose that this process makes all $\tilde{y}_j \ge 0$. Inequality $\tilde{y}_j \ge 0$ means that anti-gradient has direction into internal of the corresponding constraints. Because our problem and constrains are linear, anti-gradient, which has direction into constrains, will has this direction in any point of corresponding hyper plate (1.17). It means that this procedure will increase the amount of $y_j \ge 0$.

Example 1.4.

Find minimum of Problem

$$I = x_1 + x_2, \quad -x_1 \le 0, \quad -x_2 \le 0, \quad x_1 - 1 \le 0, \quad x_2 - 1 \le 0. \tag{1.20}$$

The equations (1.18),(1.19) are

$$\begin{aligned} -y_1 x_1 = 0, \quad y_3(x_1 - 1) = 0, \quad 1 - y_1 + y_3 - 0, \\ -y_2 x_2 = 0, \quad y_4(x_2 - 1) = 0, \quad 1 - y_2 + y_4 = 0. \end{aligned} \tag{1.21}$$

Chose equations $x_1 - 1 = 0, \quad x_2 - 1 = 0$. From solution of them we have $\tilde{x}_1 = 1, \quad \tilde{x}_2 = 1$. They satisfy (1.20). From the first column of (1.21) we get y_1 - y_2 = 0, and from the last column (1.21) we find $y_3 = y_4 = -1$. Inequality $y_i \ge 0$ is not satisfied. Change equalities by others $\tilde{x}_1 = 0, \quad \tilde{x}_2 = 0$. We get $\tilde{y}_1 = \tilde{y}_2 = 1 > 0$. Hence $\tilde{x}_1 = \tilde{x}_2 = 0$ is point of the global minimum.

Example 1.5.

Find point of global minimum in Problem

$$I = -x_1 - x_2, \quad x_1 + x_2 \le 0.$$

Solution. Write equations (1.18),(1.19)

$$y(x_1 + x_2) = 0, \quad -1 + y,$$

From $x_1 + x_2 = 0$ we get $\tilde{x}_1 = -\tilde{x}_2$. From -1+$y$ = 0 we get y = 1 > 0. Sence any $\tilde{x}_1 = -\tilde{x}_2$ is optimal.

7. Application of method β-function to quadratic programming.

This problem is following:

$$I = \sum_{j=1}^{n}\sum_{i=1}^{n} c_{ij} x_i x_j, \quad \sum a_{kj} x_j - b_k \le 0, \quad k = 1,2,..., m. \tag{1.22}$$

Assume that quadratic form in function (1.22) is positive. If don't consider constraints in (1.22), it is obvious the point of minimum in this problem is $x_j^* = 0.$ If this point satisfy inequalities in (1.22), the process of solution is finished. In particular, we have this case when all $b_k \ge 0$. We consider not triviality case. Take $\omega_j = y_j$. Equations (1.13) and (1.14) are:

$$y_k (\sum_{j=1}^{n} a_{kj} x_j - b_k) = 0 \quad i, k = 1,2,..., n; \quad \sum_{j=1}^{n} c_{ij} x_j + \sum_{j=1}^{m} y_l a_{jk} = 0, \quad y_k = 0. \tag{1.23}$$

Later procedure is analogous of the Linear Programming.

Example 1.6.

Problem are:

$$I = 0.5x_1^2 + 0.5x_2^2, \quad -x_1 - x_2 + 1 \le 0, \quad x_1 - 1 \le 0, \quad x_2 - 1 \le 0. \tag{1.24}$$

The equations (1.23)

$$\begin{aligned} &y_1(-x_2 - x_1 + 1) = 0, \quad y_2(x_1 - 1) = 0, \quad y_2(x_2 - 1) = 0 \\ &x_1 - y_1 + y_2 = 0, \quad x_2 - y_1 + y_2 = 0 \end{aligned} \tag{1.25}$$

Take the 2-nd and 3-rd equations. We get $\tilde{x}_1 = \tilde{x}_2 = 1$. The inequalities (1.24) are satisfied, but from two the last equations (1.25) for y_1 = 0 we have $\tilde{y}_2 = \tilde{y}_3 = -1$. It is contrary the request $\tilde{y}_i \ge 0$.

Take the 1-st equation in (1.25). We have $\tilde{x}_2 = 1 - \tilde{x}_1$. Solve it together with equations $\tilde{x}_1 - \tilde{y}_1 = 0, \quad \tilde{x}_2 - \tilde{y}_1 = 0$ we get $\tilde{x}_1 = \tilde{x}_2 = 1/2, \quad \tilde{y}_1 = \tilde{y}_2 = 1/2 > 0$. Hence x_1=x_2=1/2 is point of global minimum.

Appendix to #1. Proof of Theorems.

1°. <u>Proof of Theorem 1.1</u>. Proof of:

Statement 3. By subtracting the inequality $\beta(x, y) \le \beta(\bar{x}(y), y)$ from $I(x) + \beta(x, y) \ge I(\bar{x}(y)) + \beta(\bar{x}(y), y)$ we get $I(x) \ge I(\bar{x})$ over P. Statement 3 of the Theorem 1.1 is proved.

Statement 1 of the Theorem 1.1 is obvious because *X=M+P* and $I(x) \ge I(\bar{x})$ over P, we have $x^* \in M$. Statement 1 of Theorem 1.1 is proved.

Statement 2. By subtracting the inequality $J \ge \bar{J}$ from $J + I \le \bar{J} + \bar{I}$ we get $I(x) \le I(\bar{x})$ over N.

Theorem 1.1 is proved.

2°. <u>Proof</u> <u>of Theorem 1.2</u>. By adding the inequality

$I(x) + \beta(x, y) \ge I(\bar{x}(y)) + \beta(\bar{x}(y), y)$ and $-\beta(x, y) \ge -\sup_X \beta(x, y)$ over *X,* we get the estimate (1.2).

3°. <u>Proof of Theorem 1.3</u>. *Statements 1, 3*. By subtracting the inequality $\beta \le \hat{\beta}$ from $I + \beta \ge \hat{I} + \hat{\beta}$ we get $I \ge \hat{I}$ over set *P*. Statement 1 follow from this.

Statement 2. By subtracting the inequality $\beta \ge \hat{\beta}$ from $\beta - I \ge \hat{\beta} - \hat{I}$ and multiply this result by -1, we get $I \le \hat{I}$ over N, The theorem 1.3 is proved.

4°. <u>Proof of Theorem 1.7</u>. Assume x_N is point of the minimum of the objective function *J*(*x*).Therefore $J'(x_N) = 0$ because *J*(*x*) is continuously and differential, x_N is single point N_i on set *X** since this is (see Theorem 1.1')

$$I(x) + J(x) \ge I(x_N) + J(x_N) .$$

This means that $J_i(x_N) = \inf_X [I(x) + J(x)]$. The function *I*(*x*), *J*(*x*) are continuously and differential, hence $I'(x_N) + J'(x_N) = 0$. But $J'(x_N) = 0$, therefore $I'(x_N) = 0$. Theorem 1.7 is proved.

5°. <u>Proof of Theorem 1.8</u>. By subtracting the inequality $\beta \ge \beta_N$ from $\beta - I \le \beta_N - I_N$ we get $I \ge I_N$ over set *X**. The Theorem 1.8 is proved.

6°. Proof of Theorem 1.9.

Sufficiency. From "1)" of (1.9) we have

$$I + \lambda_i F_i + \omega_j \Phi_j \ge \bar{I} + \bar{\lambda}_i \bar{F}_i + \bar{\omega}_j \bar{\Phi}_j .$$

From this and "4)" (1.9) we get $I + \lambda_i F_i + \omega_j \Phi_j \ge \bar{I}$. Look it inequality over X*. On X* we have $\lambda_i F_i = 0, \quad \omega_j \Phi_j \le 0$ hence $I(x) \ge I(\bar{x})$. Because $\bar{x} \in X^*$ hence $\bar{x}$ is the point of global minimum of *I*(*x*) on *X**.

Necessity. (Method of designing). Assume that $x^* \in X$ exists. Design β (*x,y*) following way. Take $\lambda_i \equiv 0$ on X^* and take functions $\lambda_i,\ \omega_j \ge 0$ such us $J(x) > m$ *on set* $X \setminus X^*$. Then we have as the result of our design

$$J(x^*) = \inf_{x \in X^*} J(x), \quad x^* \in X^*, \quad \omega_j \geq 0, \quad \overline{\beta} = 0 .$$

The theorem 1.9 is proved.

§2. Method of combining of the extremes.

Let us to have the problems:

Problem 1 $I(x^*) = \inf\ I(x), \quad x \in X^*;$

Problem 2 $J(\bar{x}) = \inf[\, I(x) + \beta(x)], \quad x \in X$;

Problem 3 $\beta(\hat{x}) = \sup\ \beta(x), \quad x \in X.$

Assume that all points $x^*, \bar{x}, \hat{x}$ are exist.

Theorem 2.1. *Let X=X*, then for every couple* $(\bar{x}_i, \hat{x}_i)$ *which satisfy the condition* $\bar{x}_i = \hat{x}_i$ *we have*

$$\bar{x}_i = \hat{x}_i = x_i^* .$$

__*Proof*__. Let $\bar{x}_i = \hat{x}_i$ Then

$$\inf\ J(x) - \sup\ \beta(x) = J(\bar{x}_i) - \beta(\bar{x}_i) = I(\bar{x}_i) - \beta(\bar{x}_i) - \beta(\bar{x}_i) = I(\bar{x}_i) .$$

But with other side from Theorem 1.2 we have $\inf\ J(x) - \sup\ \beta(x) \leq \inf\ I$. That is $I(\bar{x}_i) \leq I(x^*)$. As x^* is point of global minimum and *X=X** hence must be only $I(\bar{x}_i) = I(x_i^*)$. As far as $\bar{x}_i$ and x_i^* exist we can find the point of minimum x_i^* such that $\bar{x}_i = x_i^*$. Theorem 2.1 is proved.

Theorem 2.2. *Let X=X*.If exist at least one of the couple* $(\bar{x}_i, \hat{x}_i)$ *such that* $\bar{x}_1 = \hat{x}_1$*, then in every point* x_i^* *we have*

1) $x_i^* = \hat{x}_i, \quad$ 2) $x_i^* = \bar{x}_i$.

__*Proof*__. 1. Assume the contrast: $\bar{x}_i \neq x_i^*$. Than summarize $I(\bar{x}_i) = I(x_i^*) \quad and \quad \beta(x_i^*) < \beta(\bar{x}_i) = \beta(\hat{x}_i)$ we get $J(x_i^*) < J(\bar{x}_i)$. This contrasts $J(\bar{x}_i) = \inf\ J(x)$.

2. Add $J(\bar{x}_i) = J(x_i^*) \quad and \quad \beta(x_i^*) = \beta(\bar{x}_i) = \beta(\hat{x}_i)$ we get $J(x_i^*) = J(\bar{x}_i)$, hence $x_i^* = \bar{x}_i$. Theorem 2.2 is proved.

From Thorems 2.1, 2.2 we have

<u>Consequence:</u>

If we want to find all points of minimum of Problem 1 it necessary and sufficiently to find all corresponding couple $(\bar{x}_i, \hat{x}_i)$.

We shall call the Problems 1 and 2 equivalents if all correspondent points of minimum of these Problems are coincided.

From Theorem 2.2 we have:

1. For equivalence of Problems 1, 2 is sufficient to exist one couple such that $\bar{x}_i = \hat{x}_i$.

2. Let exist β-functional and although one of couple $(\bar{x}_i, \hat{x}_i)$ such that $\bar{x}_i = \hat{x}_i$.

Then any points of minimum of Problem 2 and point of maximum of Problem 3 is point of minimum of Problem 1, and back, any point of minimum of Problem 1 is point of minimum of Problem 2 and point of minimum of Problem 3.

Remarks:

1. If $\beta(\bar{x}) = 0, \quad \text{then} \quad \inf J(x) = \inf I(x)$.

2. If $\bar{x} = \hat{x}$, then the lower estimate (1.1) in §1 coincide with infinum of the functional *I*(*x*).

From consequence 1 §2 we have the following

Algorithm 3. (Method of combining the extremes)

Let us take some bounded functional $\beta(x,y)$ where y is an element of the set Y. We solve this problem

$$\inf[I(x) + \beta(x,y)], \quad x \in X^*$$

and find the point of minimum

$$\bar{x}_1 = \bar{x}_1(y).$$

From

$$\sup \beta(x,y)$$

we find

$$\bar{x}_2 = \bar{x}_2(y).$$

After this we equate

$$\bar{x}_1(y) = \bar{x}_2(y) \qquad (2.1)$$

and from this equation of the combination of extreme we find the roots y_i.

These roots are the points of minimum for Problem 1:

$$\bar{x} = \bar{x}_1(y_i) = \bar{x}_2(y_i)$$

Since the Problem of finding of minimum is reduced to Problem of finding at least one root of equation of the combination of extremes (2.1).

The exist and difficulty of finding of roots dipend from chouse of β-functional, from freedom of its deformation, which give the "*y*" relation.

Note that is differ from the regular method of finding of minimum. In the usual method we take partial derivatives, equal its zero, get the set equation and from them we find only the stationary (extreme) points. They may be points local minimum, maximum, or inflection. By this method we find points of global minimum.

Thus we find the connect two various (different) problems.

The existence of solution in equation of the combination of extremes is sufficient condition for the existence of absolute minimum of functional in Problem 1.

The mathematic has good achievements in the field of existence of solution of equations. And equation (2.1) give connection between these problems and give some opportunity in solving of optimals problems.

Note also that equation (2.1) not requests that functional was continuous and differential function, hence it has wider domain for application.

If point of minimum cannot be get in explicit form than we can write this equation in form

$$\varphi_1(x, y) = 0, \quad \varphi_2(x, y) = 0\,, \tag{2.1'}$$

where function φ_1, φ_2 are got from

$$\inf_X J(x, y), \quad \sup_X \beta(x, y)\,.$$

Example 2.1. Find a point of minimum of functional

$$I = 2x^4 + x^2 - 2x + 1, \quad -\infty < x < \infty$$

<u>*Solution*</u>: Use algorithm 3. Take

$$\beta = -yx^2 + 2x\,.$$

Than

$$J = I + \beta = 2x^4 + (1 - y)x^2 + 1\,.$$

Denote x^2=w and substitute in J:

$$J = 2w^2 + (1-y)W + 1.$$

Find point of minimum this functional

$$J'_w = 4w + (1-y) = 0, \quad \overline{w} = \overline{x}_1^2 = \frac{1}{4}(y-1)$$

and point of maximum functional β:

$$\beta(x) = -yx^2 + 2x, \quad \beta'_x = -2yx + 2 = 0, \quad \overline{x}_2 = 1/y.$$

Equate $\overline{x}_1$ to $\overline{x}_2$

$$\overline{x}_1^2 = \overline{x}_2^2 \quad \frac{1}{4}(y-1) = \frac{1}{y^2}, \quad y^3 - y^2 - 4 = (y-2)(y^2 + y + 1)$$

This equation has only alone root $\overline{y} = 2$. Since $\overline{x} = \frac{1}{y} = \frac{1}{2}$.

§3. Remark about γ-functional

A) Let us take

$$\beta(x) = [\gamma(x) - 1]I(x) \tag{3.1}$$

then

$$J(x) = I(x)\gamma(x).$$

This form of common functional is sometimes more comfortable because we can chouse the multiplier to $I(x)$ which make $J(x)$ simpler.

Using our results about β-functional for this case we get following:

If X=X^* and we finding the point of global minimum Problem 2:

$$\inf_X J(x) = \inf_X [I(x)\gamma(x)] \tag{3.2}$$

than

1) Set

$$M = \{x : J - I \geq \overline{J} - \overline{I}, \quad x \in X\}$$

contains the point of global minimum of Problem 1;

2) Set

$$N = \{x:\ I\gamma + I \le \bar{I}\bar{\gamma} + \bar{I}, \quad x \in X\}$$

contains the better or same solutions than $\bar{x}$ (that is over *N*, we have $I(x) \le I(\bar{x})$);

3) Set

$$P = \{x : J - I \le \bar{J} - \bar{I}, \quad x \in X\}$$

contains the worse or same solutions than $\bar{x}$ (that is over *P*, we have $I(x) \ge I(\bar{x})$).

All these statement follow from (3.1) and Theorem 1.1.

Lower estimate (from Theorem 1.3 and (3.1) look as

$$I(x) \ge \inf_X J - \sup_X (J - I)\,. \tag{3.3}$$

Condition of equivalence of Problem 1 and 2 (theorem 2.1) in this case (*X*=*X**) is:

$\bar{x}$ and $\hat{x}$, which are founded from problems

$$\inf_X J(x) \quad \textit{and} \quad \sup_{X^8} [J(x) = I(x)]\,,$$

must equal respectively.

Algorithm 3 (Method of combining the extremes) is used for this case without change.

B) However for this case we get some new results.

Let define functional $\gamma(x,y) \ne 0$ over set $X \times Y$. We call it as <u>γ-functional</u>. Take functional

$$J(x, y) = I(x)\gamma(x, y)$$

Theorem 3.1.

Assume X=X, $\bar{x}$ is point of global minimum of Problem 2:*

$$\inf J(x), \quad x \in X, \quad \textit{where} \quad J = I(x)\gamma(x),$$

Then:

1) Set

$$P = \{x:\ 0 < \gamma \le \bar{\gamma}\}$$

contains worth or same solutions of Problem 1 (that is $I(x) \geq I(\bar{x})$ *over P);*

2) Set

$$N = \{x:\ 0 > \bar{\gamma} \geq \gamma\}$$

contains better or same solution of Problem 1 (that is $I(x) \leq I(\bar{x})$ *over N);*

3) The point of global minimum is in set $M = X \setminus \beta$, *where* $\overset{o}{P} = \{x:\ 0 < \gamma < \bar{\gamma}\}$.

Proof: 1. From inequalities $I\gamma \geq \bar{I}\bar{\gamma},\quad 0 < \gamma \leq \bar{\gamma}$ we have $I \geq \bar{I}\bar{\gamma}/\gamma,\quad \bar{\gamma}/\gamma \geq 1$. That is $I \geq \bar{I}$.

2. From inequalities $I\gamma \geq \bar{I}\bar{\gamma},\quad 0 > \gamma \geq \bar{\gamma}$ we get $I \leq \bar{I}\bar{\gamma}/\gamma,\quad \bar{\gamma}/\gamma \leq 1$. That is $I \leq \bar{I}$.

3. Because *X=M+P* and $M \cap \overset{o}{P} \neq 0$, we have $M = X - \overset{o}{P}$. Theorem is proved.

Theorem 3.2. *Assume* $\sup\limits_X \gamma > 0$. *Then we have the lower estimation*

$$I(x) = \frac{\bar{J}}{\sup \beta} \quad on \quad X\ . \tag{3.4}$$

If $\sup\limits_X \gamma(x,y) > 0 \quad for \quad \forall y \in Y$ *, we have the lower estimate*

$$I(x) \geq \sup_Y \left(\frac{\bar{J}}{\sup\limits_X X} \right). \tag{3.4'}$$

Proof: 1) For written conditions from $I\gamma \geq \bar{I}\bar{\gamma}$ we got $I \geq \bar{J}/\gamma \quad and \quad I \geq \bar{J}/\sup\limits_X \gamma$.

2) Take this estimate by *y*, we get expression (3.4)'.

Example 3.1. Find the lower estimate for functional

$$I = (x^2 - \cos x + 1)e^{(x-1)^2} \quad -\infty < x < \infty .$$

Take

$$\gamma = e^{-(x-1)^2} .$$

Then

$$J = x^2 - \cos x + 1.$$

Is it obvious the point of minimum this functional

$$\bar{x} = 0, \quad \bar{\gamma} = 1 > 0, \quad \sup_X \gamma = 1.$$

Use the estimate (3.4) we get $I(x) \geq 0$. But for $x = 0$ we have $I(0) = 0$. That way $x = 0$ is point of global minimum.

§4. Application β - function to the multi-variables nonlinear problems of constrained optimization and to problems described by regular differential equations.

A) The first problem is following. Find minimum of functional

$$I = f_o(x)\ , \tag{4.1}$$

Where *x*-*n*-dimensional vector, which satisfy independent equations

$$f_i(x) = 0, \quad i = 1,2,\ldots,m \leq n\ . \tag{4.2}$$

Functions $f(x)$ is defended in the open domain *n*-dimensional vector of space *X*. The admissible set *X** separate from *X* by equations (4.2).

Let us take some functional $\beta(x)$, such that to find

$$\inf[\ f_0(x) = \beta(x)] \quad \text{on} \quad X^*.$$

It is easier to solve.

Then from solution of Problem 2 in accordance with thorems of §1 we get the following information about Problem 1:

1) The point of global minimum is in set $M = \{x:\ \beta(x) \geq \beta(\bar{x})\}$;
2) The set $N = \{x:\ 2f_0 + \beta \leq 2\bar{f}_0 + \bar{\beta}\}$ contains better ans same solutions (that is $f_0(x) \leq f_0(\bar{x}) \quad \text{on} \quad N$);
3) The set $P = \{x:\ \beta(x) \leq \beta(\bar{x})\}$ contains worth and same solutions (that is $f_0(x) \geq f_0(\bar{x})$ on P;
4) If $X = X^* \subseteq P$, that $\bar{x}$ is point of global minimum of problem 1 (consequence 3 of §1).

Let us assume we widen the set *X** for simplification of solution. For example, we rejecte the part of constrains (4.2). Then we have

5) If $X^* \cap M = \varnothing$, than $J(\bar{x})$ is lower estimation $f_0(x)$ on *X** (consequence 5, §1).

It is more comfortable some times to take the suitable J(x) at first and find the point minimum of problem inf *J*(*x*) on *X**.

Then the corresponding sets will be (from theorem 1.1')

$$M=\{x:\ J-I\ge \bar{J}-\bar{I}\},\ \ N=\{x:\ J+I\le \bar{J}+\bar{I}\},\ \ P=\{x:\ J-I\le \bar{J}-\bar{I}\}.$$

If we solve the problem $\beta(\hat{x})=\sup\beta(x) \quad on \quad X\supseteq X^*$ we get the additional lower estimate

$$f_0(x)\ge f_0(\bar{x})+\beta(\bar{x})-\beta(\hat{x}),$$

(theorem 1.3) and set

$$M=\{x:\ f_0+\beta\le \hat{f}_0+\hat{\beta}\},\ \ N=\{x:\ \beta-f_0\ge \hat{\beta}-\hat{f}_0\},\ \ P=\{x:\ f_0+\beta\ge \hat{f}_0+\hat{\beta}\}.$$

(theorem 1.4).

Take series β_i we can get the solution of one from Problems of §1 or to facilitate thesolution of Problem 1.

The example for case *X*=X* was over (see Examples 1.1-1.3). Explain by simple examples (how you can apply the method β-functional for case, when *X*≠X* that is problem with constrains.

<u>Example 4.1</u>. Find minimum of functional

$$I=x \quad on \quad x^2+y^2-1=0\ .$$

Take any admissible point, for example $\bar{x}_0=1, \quad \bar{y}_0=0 \quad and \quad J(x)$ functional as

$$J_1=(x-x_0)^2.$$

The point of minimum of this functional is obvious $\bar{x}=x_0$. The set *M*, containing the point of global minimum, is

$$J_1-I\ge \bar{J}_1-\bar{I}, \quad that \quad is \quad (x-1)^2-x\ge -1 \quad or \quad |x-3/2|\ge 3/2$$

The boundaries of this inequality together with admissible subset (circle) draw on fig.1.3a. We see the point of absolute minimum is in left half of circle.

Take now the admissible point $\bar{x}_0=-1, \quad y=0$ and *J*-functional in more common case as

$$J_2=c(x-x_0)^2, \quad c>0.$$

Then *M* set is

$$cx^2+2cx+c-x\ge 1.$$

Take c = 0.5. Then we get $|x| \geq 1$ (fig. 1.3b).

Set *M* contain only two admissible point : x_1=1 and x_2= -1. But point x_1=1 from the J_1 cannot be the point of absolute minimum. Since the point of global minimum is $\bar{x} = -1, \quad \bar{y} = 0$.

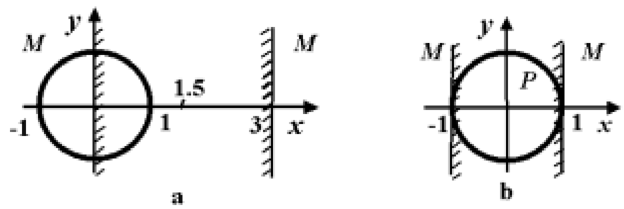


Fig. 1.3

<u>Example 4.2</u>. Find the point of global minimum of functional with constrain

$$I = x^2 - x + y^2 - 2y + 1, \quad y - \ln\left(x \pm \sqrt{x-1}\right) = 0 .$$

Take *J* functional

$$J = (x - x_0)^2 + (y - y_0)^2 .$$

The set *M* is separated by inequality

$$J - I \geq \bar{J} - \bar{I}, \quad or \quad 2y(1 - y_0) \geq (2x_0 - 1)x + a ,$$

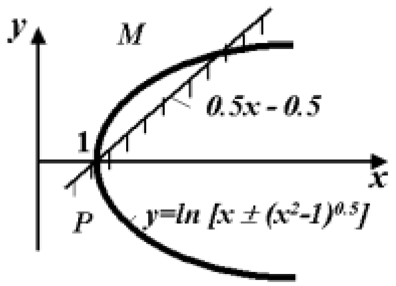
where

$$a - x_0 - 2x_0^2 + 2y_0 - 2y_0^2 .$$

Take the admissible point $x_0 = -1, \quad y_0 = 0$. Then

$$M = \left\{ x, y : y \geq \frac{1}{2}x - \frac{1}{2} \right\} \qquad \text{(Fig.1.4).}$$

From drawing we see *M* is small domain and find the point of global minimum no difficult.



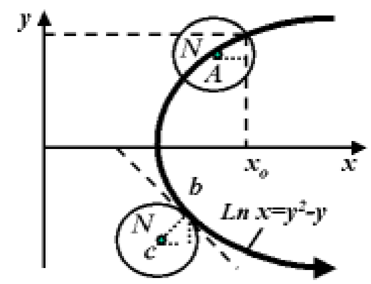


Fig. 1.4, Fig.1.5.

Example 4.3. Given functional and constrains is

$$I = 2x + 2y, \quad \ln x = y^2 = y$$

Take

$$J = (x - x_0)^2 - (y - y_0)^2,$$

where couple x_0, y_0 is admissible point.

The set *N* is separated with according Theorem 1.1 by inequality $J + I \leq \bar{J} + \bar{I}$, that is

$$\left[x - (x_0 - 1)^2\right] + \left[y + (y_0 - 1)\right]^2 \leq 2.$$

This is interior of the circle (fig.1.5).

Assume that a center of this circle is located in the point *A*. The set *N* intersect with admissible curve $\ln x = y^2 - y$. If we take a point x_0, y_0 from this intersection, we will descent along this curve whole the set *N* become by point. This take place in point *B*, where the tangent to admissible curve has the angle -45^0 (because the center of the circle is located from point x_0,y_0 from -1, -1, that is the angle $+45^0$, (fig. 1.5). Any moving from this point will return us to it.

May be shown that the point *B* is the point of global minimum.

Take into consideration when we have used the methods of β-functional (Chapter 1) we have not used in continuously and differ of functional (4.1) and constancies (4.2) unlike from known methods (for example, theory of extreme functions).

B) Consider how we can apply the methods given in §1 to optimization problems are described by regular differential equations. Below we write the statement of problem, which we widely use in future.

Assume that the moving of object is described by set of independent differential equations

$$\dot{x}_i = f_i(t,x,u), \quad i=1,2,...,n, \quad t \in T=[t_1.t_2], \tag{4.3}$$

where $x(t)$ is n - dimensional continually piece-differential vector-function of the phase coordinates, $x \in G(t)$; $u(t)$ is n - dimensional function which continuous on T except the limited number of point where it can have discontinuities of the 1-st form, $u \in U$ is an independed variable. Boundary values t_1, t_2 is given, $x(t_1) \in G(t_1)$, $x(t_2) \in G(t_2)$.

The aim function is

$$I = F(x_1,x_2) + \int_{t_1}^{t_2} f_0(t,x,u)dt, \quad x_1 = x(t_1), \quad x_2 = x(t_2). \tag{4.4}$$

Functions $F(x_1,x_2)$, $f_i(t,x,u)$, $i = 0,1,\ldots,n$ are continuous over $T \times G \times U$. Set of continuous, almost

everywhere differentiable functions $x(t) \in G(t)$ we denote D. Set of pies-continuous functions $x(t) \in U$, we denote V. Set of couple $x(t)$, $u(t)$ which satisfy these requirements and almost everywhere comply with equations (4.3) we shall call **admissible** and denote Q, $Q \subset D \times V$.

Consider the problems:

a) Find the coiple $u^*(t)$, $x^*(t) \in D$, which give the minimum of function (4.4) (Traditional statement).
b) Find sup-set $N \subset G \times U \times T$ such that any admissible curve from N we have $I(x) \le c$, where c is constant.
c) Find the lower estimate of $I(x)$ over Q.

Take the function $\int_{t_1}^{t_2} \beta(t,x,u)dt$, where $\beta(t,x,u)$ is a definite and continuous function on $T \times G \times U$.

Theorem 4.1. *Let us assume that $F \equiv 0$ and Problem 2 is solved. That means*

$$J(\bar{x},\bar{u}) = \inf J(x,u) \quad \text{on} \quad Q,$$

where

$$J = \int_{t_1}^{t_2} [f_0(t,x,u) + \beta(t,x,u)]dt.$$

Then:

1) Set

$$N = \{t,x,u: \ 2f_0 + \beta \le 2\bar{f}_0 + \overline{\beta}, \quad t \in T\}$$

contains the same or better solutions of Problem 1.

3) Set

$$P = \{t,x,u: \ \beta \le \overline{\beta}, \quad t \in T\}$$

contains the same or worse solutions of Problem 1.

Proof: 1. On set Q from N we have

$$\int_{t_1}^{t_2}(2f_0+\beta)dt \le \int_{t_1}^{t_2}(2\bar{f}_0+\bar{\beta})dt .$$

Subtract from this inequality following

$$\int_{t_1}^{t_2}(f_0+\beta)dt \ge \int_{t_1}^{t_2}(\bar{f}_0+\bar{\beta})dt , \qquad (4.5)$$

we get over *Q* from *N*

$$\int_{t_1}^{t_2} f_0 dt \le \int_{t_1}^{t_2} \bar{f}_0 dt .$$

2. By analogy with above, subtract from inequality

$$\int_{t_1}^{t_2}\beta dt \le \int_{t_1}^{t_2}\bar{\beta} dt$$

the inequality (4.5) we get over Q from *P*

$$\int_{t_1}^{t_2} f_0 dt \ge \int_{t_1}^{t_2} \bar{f}_0 dt .$$

The Theorem 4.1 is proved.

Sets *N*, *P* not empty. They contain at least one curve from Q. This curve is $\bar{x}(t), \bar{u}(t) \in Q$.

If we solve the additional problem

$$\sup_{Q} \int_{t_1}^{t_2}\beta dt ,$$

we get additional information about sets *N*, *P* and lower estimate. It is following

Theorem 4.2. *Let us assume F ≡0 and solved the Problem*

$$\sup \int_{t_1}^{t_2}\beta(t,x,u)dt \quad \text{on} \quad Q .$$

Then

1) *Set*

$$N = \{t,x,u: \ \beta - f_0 \ge \hat{\beta} - \hat{f}_0 , \quad t \in T\}$$

contains the same or better solutions:

2) Set

$$P=\{t,x,u:\ f_0+\beta\le\hat{f}+\hat{\beta},\quad t\in T\}$$

contains the same or worse solutions.

Here $\hat{f}_0=f_0(t,\hat{x},\hat{u}),\quad \hat{x}(t),\quad \hat{u}(t)$ is curve of extreme

$$\sup\int_{t_1}^{t_2}\beta(t)\quad \text{on}\quad Q\,.$$

Proof: 1. Over *Q* from *N* we have

$$\int_{t_1}^{t_2}(\beta-f_0)dt\ge\int_{t_1}^{t_2}(\hat{\beta}-\hat{f}_0)dt$$

Subtract from this inequality the following

$$\int_{t_1}^{t_2}\beta dt\le\int_{t_1}^{t_2}\hat{\beta}dt\,,$$

we get

$$\int_{t_1}^{t_2}f_0dt\le\int_{t_1}^{t_2}\hat{\hat{f}}_0dt\,.$$

2. By analogy, subtract $\int_{t_1}^{t_2}\beta dt\le\int_{t_1}^{t_2}\hat{\hat{\beta}}dt$ from

$$\int_{t_1}^{t_2}(f_0+\beta)dt\ge\int_{t_1}^{t_2}(\hat{\hat{f}}_0+\hat{\hat{\beta}})dt$$

we get

$$\int_{t_1}^{t_2}f_0dt\ge\int_{t_1}^{t_2}\hat{\hat{f}}_0dt\,.$$

The Theorem 4.2 is proved.

Theorem 4.3. *(Lower estimation).*

Assume F $\equiv$0, the ends of x(t) are fixed, β (t,x,u) is defined and bounded on G$\times$U$\times$T.

Then there is lower estimate of Problem 1:

$$I(x,u)\ge\int_T[\,f_0(t,\overline{x},\overline{u})+\beta(t,\overline{x},\overline{u})-\beta(t,\hat{x},\hat{u})]dt \tag{4.6}$$

Proof: Subtract $\int_T \beta dt \le \int_T \sup \beta dt$ from inequality

$$\int_T (f_0 + \beta)dt \ge \int_T (\bar{f}_0 + \overline{\beta})dt$$

we get (4.6). The theorem 4.3 is proved.

Consequence 1: Couple $\bar{x}, \bar{u}$ is curve of absolute minimum of Problem 1 over set *N*.

Consequence 2: If set P⊇T×G×U (or accessible) than $\bar{x}, \bar{u}$ (or $\hat{x}, \hat{u}$) is curve of global minimum of problem 1 over *Q*.

Similar results we can get for case, when $F \neq 0$ and ends of *x(t)* can move.

Example 4.4. Assume the problem is described by conditions:

$$I = \int_0^1 (x^2 + e^u)dt, \quad \dot{x} = u, \quad |u| \le 1, \quad x(0) = 1, \quad x(1) = 0.$$

Use the theorem 4.1. Take $\beta = -e^{+u}$. We get the problem

$$I = \int_0^1 x^2 dt, \quad \dot{x} = u, \quad |u| \le 1, \quad x(0) = 1, \quad x(1) = 0.$$

Its solution is $\bar{x} = -t, \quad \bar{u} = -1, \quad 0 \le t \le 1.$

Find set *P*: $\beta \le \overline{\beta}. \quad That \quad is \quad e^u \ge e^{-1}, \quad u \ge 1$.

But value *u* < -1 is not acceptable. Since *P* is cover all admissible set points *t,x,u*. That way $\bar{x} = -t$.

Is the curve of global minimum (see Consequence 2).

Example 4.5. Find of minimum in problem

$$I = \int_0^2 (|x| + 0.5x^2)dt, \quad \dot{x} = u, \quad |u| \le 1, \quad x(0) = 1, \quad x(2) = 0.$$

We have here undifferentiated function in integral. Known methods us variational calculation or principle of maximum are not been used.

Change this problem following "good" (easy) problem:

$$I = \int_0^2 0.5x^2 dt, \quad \dot{x} = u, \quad |u| \le 1, \quad x(0) = 1, \quad x(2) = 0$$

and find

$$\sup_{x(t)} L .$$

The solution is shown in Fig. 1.6.

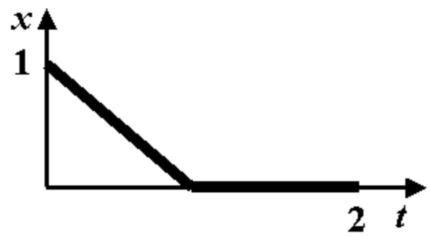


Fig. 1.6. For Example 4.5.

By according the theorem 4.2

$$P = \{x : |x| \geq |\bar{x}|\},$$

that means set *P* cover all accessible domain. Since abtained, solution is curve of global minimum of Problem 1.

5. Method of β - function in minimizing sequences

A) The sequence {x_s} such that $I(x_s) \underset{s\to\infty}{\to} \inf I(x)$ on the set *X** is named as a minimizing sequence (for Problem 1).

We must design these sequence in a successive approximation methods and in case, when extreme is absent in an allowable (admissible) subset.

Theorem 5.1. *Assume β (x) ≤ 0 on X* and there exist sequence {x_s}$\in$X* such, that*

$$J(x_s) \to \inf J \quad \text{for} \quad s \to \infty \quad \text{on} \quad X \tag{5.1}$$

Then: 1) $I(x_s) \underset{s\to\infty}{\to} m = \inf I(x) \quad on \quad X^*$;

2) Any sequence {x_s}$\in$*X, which satisfy* (5.1) *or* $I(x_s) \to \inf_X J$ *, minimize I(x) on X*, minimize and J(x) on X.*

Proof: 1. Because β *(x)* ≤ 0 on *X**, we have $\inf_X J \leq I(x)$. That is $\inf_X J \leq \inf_{X^*} J$. From {$x_s$}$\in$*X** and (5.1) we have that

$$\inf_X J = \inf_{X^*} I \,. \tag{5.2}$$

That is $I(x_s) \to m$.

2. From (5.1) and (5.2) we have the statement 2 of the theorem.

3. From $I(x_s) \underset{s\to\infty}{\to} m$ and (5.2) we have $J(x_s) \to \inf J \quad \text{for} \quad s \to \infty \text{ on } X$. Theorem is proved.

Remark. The requirement $\beta(x) \le 0$ on X^* of the theorem 5.1 we can change by the requirement $\sup_{X^*} \beta \le 0 \quad \text{on} \quad X^*$ because from sup $\beta \le 0$ on X^* we have $\beta(x) \le 0$ on X^*.

Theorem 5.2. *Assume there exist the sequence $\{x_s\} \in X^*$ such that*

$$J(x_s) \underset{s\to\infty}{\to} \inf J(x) \ \text{ on } X \text{ (or } X^*) \quad \text{and} \quad \beta(x_s) \to \sup \beta \quad \text{on } X \text{ (or } X^*) \tag{5.3}$$

Then this sequence is minimized.

<u>Proof</u>*:* From $I(x_s) + \beta(x_s) \to \inf J$ and $\beta(x_s) \to \sup \beta$ we get that $I(x_s) \to \inf J - \sup \beta$. Because
$I(x_s) \ge \inf J - \sup \beta$ *and there exist* $\{x_s\} \in X^*$ *we have* $I(x_s) \to m = \inf J - \sup \beta$.

Q.E.D.

Remark: From (1.1) and (1.1') we see that *X* and X* in (5.3) we can take in any combinations.

B) Let us consider a case now, when we have both a sequence of elements $\{x_s\}$ and a sequence of functions $\{\beta_i(x)\}$.

Theorem 5.3. *In order that a sequence* $\{x_s\} \in X^*$ *minimize function I(x) on set X*. It is sufficient that there exist a sequence of functions $\{\beta_i(x)\}$ such that*

(1) $\beta_i(x) \le 0$ over X^* *for all i*;

(2) *There exist numbers* $q_i = \inf_X J_i$, $q = \lim_{i\to\infty} q_i$;

(3) $J(x_s) \to q$ *or* $I(x_s) \to q$ *if* $s \to \infty$.

This theorem may be proved easy, because $q = \inf I$ over set X^*.

From theorems 2.1, 2.3 we have next *statement*:

If there exist one sequence which satisfy theorem 2.3 then any other sequence which belong to set *X*, $\{x_s\} \in X$ and satisfy the condition $I(x_s) \to q$ or $J(x_s) \to q$ is minimize for Problem 1.

Appendix to Chapter 1.

1. **Operations with signs** inf **and** sup**.**

Below there shown the characteristics of signs **inf** and **sup,** which can be useful for solution of problems. The proof is simply and no given. We assume that are shown constrains have place in domain of definition of function.

1. $\inf[-f(x)] = -\sup f(x), \quad \sup[-f(x)] = -\inf f(x).$

2. $\inf cf(x) = c\inf f(x) \quad \text{if} \quad c = const > 0;$

$\inf cf(x) = -c\inf f(x) \quad \text{if} \quad c = const < 0.$

3. $\inf[c + f(x)] = c + \inf f(x),$

4. $\inf \dfrac{1}{f(x)} = \dfrac{1}{\sup(x)} \quad \text{if} \quad f(x) \neq 0.$

5. If $\bar{x}(t)$ can have breaks and $f(t, \bar{x}(t))$ has integrality then

$$\inf_{x(t)} \int_{t_1}^{t_2} f[t, x(t)]dt = \int_{t_1}^{t_2} \inf_{x} f(t, x)dt .$$

6. Assume *f(φ)* is monotone function, $\partial f / \partial \varphi$ is continuous. Then

$$\inf f[\varphi(x)] = f[\inf_X \varphi(x)] \quad if \quad \partial f / \partial \varphi > 0,$$

$$\inf f[\varphi(x)] = f[\sup_X \varphi(x)] \quad if \quad \partial f / \partial \varphi < 0 .$$

Consequences

Functions of single variable

$a)$ $\inf f^{2n}(x) = [\inf f(x)]^{2n}$, if $f(x) \geq 0$,

$\inf f^{2n}(x) = [\inf f(x)]^{2n}$, if $f^{2n-1}(x) > 0$,

$\inf f^{2n}(x) = [\sup f(x)]^{2n}$, if $f^{2n-1}(x) < 0$.

$b)$ $\inf f^{2n+1}(x) = [\inf f(x)]^{2n+1}$,

$c)$ $\inf \log_a f(x) = \log_a \inf f(x)$, if $a > 1$.

$\inf \log_a f(x) = \log_a \sup f(x)$, if $0 < a < 1$.

$d)$ $\inf a^{f(x)} = a^{\inf f(x)}$, if $a > 1$.

$\inf a^{f(x)} = a^{\sup f(x)}$, if $0 < a < 1$.

$e)$ $\inf \sin f(x) = \sin \inf f(x)$ in domain $(-0.5\pi \leq x \leq 0.5\pi)$.

$f)$ $\inf \cos f(x) = \cos \sup f(x)$ in domain $(0 \leq x \leq \pi)$.

$g)$ $\inf a \tan f(x) = a \tan \inf f(x)$.

$h)$ $\inf \tan f(x) = \tan \inf f(x)$, if $|f(x) < \pi/2|$.

$i)$ $\inf \sqrt{f(x)} = \sqrt{\inf f(x)}$.

$j)$ $\inf\limits_t \frac{df(x)}{dt} = \frac{d}{dt} f(x)\Big|_{\bar{t}}$ in domain $f''(t) > 0$. Here $\bar{t} = \arg \inf\limits_t f(t)$.

$k)$ $\inf\limits_t \frac{df(x)}{dt} = \frac{d}{dt} f(x)\Big|_{\bar{t}}$ in domain $f''(t) < 0$. Here $\bar{t} = \arg \sup\limits_t f(t)$.

Estimates

A. Functions of single variable

1. $\inf[f_1(x) + f_2(x)] \geq \inf f_1(x) + \inf f_2(x)$.
2. $\inf[f_1(x) f_2(x)] \geq \inf f_1(x) \inf f_2(x)$ if $f_1(x) > 0$, $f_2(x) > 0$.
3. $\inf \frac{f_1(x)}{f_2(x)} \geq \frac{\inf f_1(x)}{\sup f_2(x)}$, if $f_1(x) > 0$, $f_2(x) > 0$.

We gave above in $1-3$ the sign $=$ if $\bar{x}_1 = \bar{x}_2$.

4. $\inf\limits_{x(t) \in Q} \int_{t_1}^{t_2} f(t, x(t))dt \geq \int_{t_1}^{t_2} \inf\limits_x f(t, x)dt$.

B. Function of two variables

1. 1. $\inf[f_1(x)+f_2(y)] = \inf\limits_x f_1(x) + \inf\limits_y f_2(y).$

2. $\inf[f_1(x)f_2(y)] \ge \inf\limits_x f_1(x) \inf\limits_y f_2(y)$ if $f_1(x)\ge 0,\ f_2(y)\ge 0.$

3. $\inf\limits_{x,y} \dfrac{f_1(x)}{f_2(x)} \ge \dfrac{\inf f_1(x)}{\sup f_2(y)},$ if $f_1(x)\ge 0,\ f_2(y) > 0.$

4. $\inf\limits_{x,y} f(x,y) f(x,y) = \inf\limits_y f(\bar{x}(y), y).$

2. Exercises for β- and γ- functions

Choosing β- function, find quasi-optimal solutions to precision 5%.

Indication: Find the lower estimate. Separate subset which contains points of global minimum and take quasi-0ptimal solution from it.

Examples: Answers:

1. $I = x^4 + x^2 + 0.2x + 1,\quad M = \{x:\ -0.2 \le x \le 0\},\ I(0) = 1 \ge 0.99.$
2. $I = x^6 + x^2 + 0.2x + 1,\quad M = \{x:\ -0.2 \le x \le 0\},\ I(0) = 1 \ge 0.99.$
3. $I = x^8 + x^2 - 0.2x + 1,\quad M = \{x:\ 0 \le x \le 0.2\},\ I(0) = 1 \ge 0.99.$
4. $I = x^{2n} + x^2 - 0.2x + 1,\quad M = \{x:\ 0 \le x \le 0.2\},\ I(0) = 1 \ge 0.99.$
5. $I = |x|^m + x^2 - 0.4x + 1,\quad M = \{x:\ -0.2 \le x \le 0\},\ I(0) = 1 \ge 0.99.$
6. $I = |x|^m + 2x^2 - x + 3,\quad M = \{x:\ -0.5 \le x \le 0\},\ I(0) = 3 > 2\,{}^7\!/_8.$
7. $I = x^2 - 4x + 6 - 0.1\sqrt[3]{e^{-(x-1)^2}},\quad M_1 = \{x:\ 0 \le x \le 2\},\ I(2) = 2 - 0.1e^{-2/3} \ge 1.9.$
   $M_2 = \{x:\ 1 \le x \le 3\}.$
8. $I = x^2 - 4x + 6 - \dfrac{0.1}{(x-1)^2 + 10},\quad M = \{x:\ 1 \le x \le 3\},\quad I(2) = 2 - \dfrac{1}{10} \ge 1.99.$
9. $I = x^2 - 2x + 5 - \dfrac{1}{x^2 - 4x + 14},\quad M = \{x:\ 1 \le x \le 2\},\quad I(1) = 4 - \dfrac{1}{11} \ge 3.9.$
10. $I = x^2 + 4x + 6 - \dfrac{0.1}{x^4 + 3x^3 + 3x^2 + 2},\quad M = \{x:\ -3 \le x \le -1\},\quad I(-2) = 1.95 \ge 1.9.$
11. $I = x^2 + 2x + 3 - \dfrac{0.1}{e^{x^2 - 2x + 1}},\quad M = \{x:\ -3 \le x \le 1\},\quad I(-1) = 2 - \dfrac{0.1}{e^2 + 1} \ge 1.9.$

12. $I = |x-1|^3 + 5 - \dfrac{0.2}{(x-1)^2 + 10}$, $M = \{x = 1\}$, $I(1) = 4.98 \ge 4.98$.

13. $I = \sqrt{x^2 - 4x + 8} - \dfrac{0.1}{(x-1)^2 + 5}$, $M = \{x:\ 1 \le x \le 3\}$, $I(2) = 2 - \dfrac{0.1}{6} \ge 2 - \dfrac{0.1}{5}$.

14. $I = \sqrt[3]{x^2 - 4x + 12} - \dfrac{0.1}{\sqrt{x^2 - 2x + 5}}$, $M = \{x:\ 1 \le x \le 3\}$, $I(2) = 2 - \dfrac{0.1}{5} \ge 2 - \dfrac{0.1}{2}$

15. $I = \sqrt{x^2 + 4x + 8} - \dfrac{0.1}{\sqrt[3]{x^4 + 3x^3 + 3x^2 + 2}}$, $M = \{x:\ -3 \le x \le -1\}$, $I(-2) = 1.95 \ge 1.9$

16. $I = |x - x_1|^n + c_1 - \dfrac{d}{|x - x_2|^m + c_2}$, $d > 0,\ c_2 > 0,\ n > 0,\ m > 0,$

$M = \{x:\ |x - x_2| \le |x_1 - x_2|\}$, $I(x) \ge c_1 - \dfrac{d}{c_2}$.

17. $I = \sqrt[k_1]{|x - x_1|^n + c_1} - \dfrac{d}{\sqrt[k_2]{|x - x_2|^m + c_2}}$, $d > 0,\ c_1 > 0,\ c_2 > 0,\ n > 0,\ m > 0,\ k_1 > 0,\ k_2 > 0,$

$M = \{x:\ |x - x_2| \le |x_1 - x_2|\}$, $I(x) \ge \sqrt[k_1]{c_1} - \dfrac{c_1}{\sqrt[k_2]{c_2}}$.

18. $I = |x(x-2)| - \dfrac{0.1}{|x-1| + 10}$, $M = \{x:\ 0 \le x \le 2\}$, $I(0) = 0 - \dfrac{0.1}{11} \ge -\dfrac{0.1}{10}$.

19. $I = |x(x-a)| - \dfrac{d}{|x-b| + c}$, $d > 0,\ c > 0,$

$M_1 = \{x:\ |x - b| \le |a - b|\}$, $M_2 = \{x:\ 0 \le x \le 2b\}$, $I \ge -\dfrac{d}{c}$.

20. $I = \dfrac{|x|}{x} + |x|.$ $M = \{x:\ x < 0\}$, $I(0) = -1 \ge -1.$ *Indication*: $\beta = -[x]$. $I(0) = -1 \ge -1$.

21. $I = x^2 - 1.8x + 1 + \dfrac{1}{1 + e^{\sin x}}$. *Answer* $M = \{x:\ 0.8 \le x \le 1\}$, $I(1+0) = 0.2 - 0 \ge 0.1$.

22. $I = x^2 - 0.2x + 10.1 + \dfrac{1}{10 + \lg|\cos x|}$, *Ansfer* $M = \{x:\ |\cos x| \ge 10.10\}$, $I(0) = 10.1 \ge 10.1$

23. $I = x^2 + xe^{4x} + 10$. *Ans.* $M = \{x : x \le 0\}$, $I(0) = 10 \ge 10 - \dfrac{1}{4e}$.

24. $I = \dfrac{e^x}{x} + x\ln^2 x, \quad x \ge 0$. *Ans.* $M = \{x : 0 < x \le 1\}$, $I(0) = e \ge e$.

25. $I = x^6 + y^6 + 2x^2 - 4xy + 2y^2$. *Ans.* $M = \{x, y : x = y\}$, $I(0, 0) = 0 \ge 0$. $I =$

26. $I = |x| - e^{-y^2} + x^2 - 2xy + y^2$. *Ans.* $M = \{x, y : \ x = y\}$, $I(0) = 1 \ge 1$.

27. $I = |x + |y - 1|| + |z + 1| + \dfrac{1}{e^{x^2+y^2+z^2}} + 6$. *Ans.* $M = \{xy, z : x^2 + y^2 + z^2 \le 2\}$, $I(0,1,-1) = 6 + \dfrac{1}{e^2} \ge 6$.

28. Find the minimum from all integer solutions of function

$$I = (x - 32)^2 + \frac{(\log_2 x - 5)(\log_2 x - 5.1)}{\lg x}.$$

Indication. Take as *β* the second member in *I* and consider the in the extended area $0 < x < \infty$. We find $M = \{x : 32 \le x \le 34.3\}$. Calculate *I* for *x* = 32, 33, 34 and select better.

Find the lower estimation by using the ***γ* –** function.

29. $I = \dfrac{x^2(x-2)^2}{2 - \sin x}$. *Ans.* $I(x) = 0$, $x_2^* = 0$, $x_2^* = 2$.

30. $I = (x-2)^2(1 + \lg^2 x)$. *Ans.* $I(x) \ge 0$, $x^* = e$.

31. $I = (|x| + |y|)e^{-(x^2+y^2)}$. *Ans.* $M = \{0, 0\}$, $I(0, 0) = 0 \ge 0$.

References to Chapter 1

1. Bolonkin A.A., A certain method of solving optimal problems, Processing Siberian Department of USSR Science Academia (Izvesiya Sibirsk. Otdel. Akadem.Nauk SSSR), 1970, No.8, issue 2, June, pp.86-92 (in Russian). Math. Reviews, v.45, #6163 (English).
2. Bolonkin A.A., New Methods of Optimization and their Application, Moscow, MVTU, 1972, 220 ps. (Russian). Болонкин А.А., **Новые методы оптимизации и их применение**. МВТУ им. Баумана, 1972г., 220 стр. (См. РГБ, Российская Государственная Библиотека, Ф-861-83/1809-6**).**
   http://vixra.org/abs/1504.0011 **v4.** http://viXra.org/abs/1501.0228**,** (v1, old) ,
   http://viXra.org/abs/1502.0137 **v3;** http://viXra.org/abs/1502.0055 v2;
   https://archive.org/details/BookOptimization3InRussianInWord20032415 v2,
   https://archive.org/details/BookOptimization3InRussianInWord20032415_201502 v3,
   https://archive.org/details/BookOptimizationInRussian **(old),**
   http://intellectualarchive.com/ **(v2) again 2 6 15 ?, v3 2 16 15, (v.4) 3 17 15 (must be<7Mb)**
   http://www.twirpx.com/file/1592607/ 2 6 15 загрузил v2.

http://www.twirpx.com/file/1605604/?mode=submit v3, загрузил 2 16 15
https://www.academia.edu/11054777/ v.4.

3. Болонкин А.А., **Новые методы оптимизации и их примеение в задачах динамики управляемых систем.** Автореферат диссертации на соискание ученой степени доктора технических наук. Москва, ЛПИ, 1971г., 28 стр. http://viXra.org/abs/1503.0081, 3 11 15. http://www.twirpx.com , https://archive.org ?(не загружается?>7 Mb?) http://samlib.ru/editors/b/bolonkin_a_a/ , http://intellectualarchive.com/ .#1488. https://independent.academia.edu/AlexanderBolonkin/Papers,
4. **Докторская диссертация** А.Болонкина: Новые метоы оптимизации и их применение в задачах динамики управляемых систем. ЛПИ 1969г. https://drive.google.com/file/d/0BzlCj79-4Dz9YTJOUHVhR1FZUVE/view?usp=drive_web Dissertation Optimization 1-2 9 30 15.doc, http://viXra.org/abs/1511.0214 , http://viXra.org/abs/1509.0267 Part 1, http://vixra.org/abs/1509.0265 Part2, https://archive.org/details/NewMethodsOfOptimizationAndItsApplication.Part1inRussian https://www.academia.edu/s/2a5a6f9321?source=link, http://www.twirpx.com,
5. Болонкин А. А., **Об одном методе решения оптимальных задач**. Известия СО Академии наук СССР, вып.2, № 8, июнь 1970 г. http://www.twirpx.com/file/1837179/ , http://viXra.org/abs/1512.0357 , http://vixra.org/pdf/1512.0357v1.pdf , https://www.academia.edu , https://archive.org/download/ArticleMethodSolutionOfOptimalProblemsByBolonkin
6. **List #1 Bolonkin's publications in 1965-1972**.(in Russian). https://archive.org/details/No1119651972
7. Bolonkin A.A., A New Approach to Finding a Global Optimum, Vol.1, 1991, The Bnai Zion Scientists Division, New York (English).
8. Bolonkin A.A., **Impulse solutions in optimization problems** 11 20 15 http://viXra.org/abs/1511.0189; https://www.academia.edu/s/00538971c8 , https://archive.org/details/ArticleImpulseSolutionsdoc200311115AfterJoseph , GSJornal: http://gsjournal.net/Science-Journals/Research%20Papers-Astrophysics/Download/6259, www.IntellectualArchive.com, #1625, https://www.researchgate.net ???

Chapter 2

Methods of *α* – functions. Estimtions.

§1. *α* – functions over arbitrary set.

A. The special case of *β*-function is α-function. It is defined over set *Z=X×Y* and has the following properties:

1) There exist subset $K \subset Z$ with projection *K* on X_i $pr_1K = X^*$.

2) $\tilde{\alpha}(x,y)=0$ on *K*.

Theorem 1.1. *Assume* $\tilde{\alpha}(x,y)$ *is* $\tilde{\alpha}$ *-function and exist the point of global minimum* $x^* \in X^*$.
Then the element $\bar{x}$ *is point of the global minimum of object function I(x) over set X* if and only if there exist* $\tilde{\alpha}(x,y)$ *such that:*

1) $J(\bar{x},\bar{y}) = \inf\,[I(x)+\alpha(x,y)] \quad x,y \in Z; \quad 2)\ \bar{x},\bar{y} \in K$.

Proof: As $\bar{x},\bar{y} \in K$, then $\alpha(\bar{x},\bar{y})=0$ and

$$J(\bar{x},\bar{y}) = \inf_Z [I(x)+\tilde{\alpha}(x,y)] = \inf_K [I(x)+\alpha(x,y)] = \inf_{X^*} I(x)\,.$$

Q.E.D.

One may made vice versa. Define set $K_1 = \{x,y : \tilde{\alpha}(x,y)=0, \quad x \in X, \quad y \in Y\}$. Find $X_1 = pr_1K_1$. Then $\bar{x}$ is the point of minimum *I(x)* over X_1, if $\bar{x},\bar{y} \in K_1$.

The special case of $\tilde{\alpha}$ -function is *α*-function defined over *Z* and such that *α(x,y)* = 0 over X* for all $y \in Y$.

The following theorem is important:

Theorem 1.2. *Let us assume α(x,y) = 0 over* X^* *for all* $y \in Y$ *and there exist* $x^* \in X^*$.
The element $\bar{x}$ *will be the point of global minimum of objective function I(x) over X* if there exist function α (x,y) such that*

1) $J(\bar{x},y) = \inf\,[I(x)+\alpha(x,y)] \quad x,y \in Z; \quad 2)\ \bar{x} \in X^*$. (1.1)

Proof: As $\bar{x} \in X^*$, then $\alpha(x,y)=0$ and

$$J(\bar{x},\bar{y}) = \inf_Z [I(x)+\alpha(x,y)] = \inf_X [I(x)+\alpha(\bar{x},y)] = \inf_{X^*} I(x)\,.$$

Q.E.D.

If *y* is not constant, one can use it (the function $\alpha(x,y)$ from *y*) for getting $\bar{x} \in X^*$.

Theorem 1.3. $\tilde{\alpha}$ *and α – functions exist and their number is infinite.*

Theotem 1.4. *(Estimate). If in (1,1)* $\bar{x} \notin X^*$ *, we have a lower estimation of the objective function I(x) on* X^**:*

$$J(\bar{x}(y), y) \le I(x) \quad for \quad all \quad y \in Y .$$

One can get this estimation from $\alpha(x, y) = 0$ on set X^* for all $y \in Y$ and **Principle of Extension**[1] [5], because $X^* \subseteq X$.

----

1) The Principle of extension state: any extension of set, which you find on a minimum of functional, can only decrease on a minimum of an objective function (can only decrease value of a minimum).

The dependence *J(x,y)* from y one may use for improving of estimation. In particular, one can take *α = α(x)*. Then from theorems 1.2, 1.3 one can get the following consequences:

<u>Consequence 1.</u> Assume *α(x)* = 0 on X^* and exist $x^* \in X^*$. Element $\bar{x}$ is point of a minimum of the objective function *I(x)* on *X** if and only if the exist α(x) such, that

1) $J(\bar{x}) = \inf\,[I(x) + \alpha(x, y)] \quad x \in X; \quad 2)\ \bar{x} \in X^*$. (1.1')

<u>Consequence 2</u>. If $\bar{x} \in X^*, \quad \beta \equiv \alpha \quad then \quad \inf\limits_{X \times Y} J = \inf\limits_{X^*} I$.

As far as *α*-function is the particular case *β*-function consequently the theorem 1.1 of Chapter 1 is right in this case.

Theorem 1.5. Assume $\bar{x}$ is point of global minimum of Problem 2:

$$J(\bar{x}) = \inf[I(x) + \alpha(x)], \quad x \in X .$$

Then: 1) The points of global minimum of Problem 1 are in the set

$$M^* = M \cap X^*, \quad where \quad M = \{x : \alpha \ge \bar{\alpha}\} ;$$

2) Set $N^* = N \cap X^*, \quad where \quad N = \{x : J + I \le \bar{J} + \bar{I}\}$, contain same or better solution that is in *N* the object function $I(x) \le I(\bar{x})$;

3) Set $P^* P \cap X^*, \quad where \quad P = \{x : \alpha \le \bar{\alpha}\}$ contains same or worse solutions (that is $I(x) \ge I(\bar{x})$ in P).

The same way for this case we can be formulated the Theorem 1.1

Since the set X^* is selected by equal $\alpha(x) = 0$ we get from Theorem 1.5 the consequences:

Consequence 3: If

$$Consequence\ 3: \quad If \quad \alpha(\bar{x}) > 0\,, \ then \quad X^* \subseteq P.$$

$Consequence\ 4: \quad If\ \alpha(\bar{x}) < 0\ ,\ then \quad X^* \subseteq M.$

$Consequence\ 5: \quad If\ \alpha(\bar{x}) = 0\ ,\ then \quad \bar{x} \in X^*.$

From Theorems 1.2 – 1.4 and Consequence 1 we get **Algorithm 4** :

We take the bounded of below functional (objective function) defined on X*Y, find minimal $\bar{x} = \bar{x}(y)$ of Problem 2: $\inf\ (I+\alpha)$, $x \in X$ or minimal in implicit form $\xi(\bar{x}, y) = 0$. We solve together the system equations (combining equations of α- function): $\inf(I+\alpha)$, $x \in X$. Then value $\bar{x}$ - root pf this system is the absolute minimal of Problem1: $\inf(I+\alpha)$, $x \in X$.

Algorithm 4' (solution by choice of α-function).

We take the bounded of below functional α defined on *X* (or *X*Y*), Solve the Problem 2: $\inf\ (I+\alpha)$, $x \in X$. If $\bar{x} \in X^*$, we get minimal of Problem 1, if $\bar{x} \notin X^*$, we get the estimation below $J(x) \le I(x^*)$ of value of the objective function *I(x)* on set X^* and we get the sets *M, N, P*.

Comments: 1. If the admissible set X^* allocates by functional $F_i(x) = 0$, you can find the α - functional in form $\alpha = \lambda_i(x)F_i(x)$ (here *i* means sum), where $\lambda_i(x)$ are some function of *x*.
2. If the admissible set allocate by functional $\Phi_j(x) \le 0$, you can find α – functional in form

$$\alpha = \omega_l(x)[\Phi_l(x) + |\Phi_l(x)|],$$

where $\omega_l(x)$ are some function of *x* , or in form

$$\alpha = \omega_l(x)\Phi_l(x),$$

where $\omega(x) \ge 0$ and it is fulfilled the condition $\omega_l(x)\Phi_l(x) \equiv 0$ on X^*.

3. Assume there is some *α* –functional and element $x \in X^*$ such $J(\bar{x}) = \inf[I(x) + \alpha(x)], \quad x \in X$. Then any element $x_1 \in X^*$ and is satisfying the condition

$$J(x_1) = \inf[I(x) + \alpha(x)], \quad x \in X\ . \qquad (1.1'')$$

is point of the absolute minimum the functional *I(x)* on X^* and any point of the absolute minimum the functional *I(x)* on X^* satisfy the condition (1.1”).

This direct statement follows immediately from condition 1.

We proof the converse. Since the global minimal $x_1 \in X^*$, it means $\alpha(x_1) = 0$, then

$$I(x_1) = \inf_{X^*} I(x) = J(x_1) = J(\bar{x}) = \inf_X [I(x) + \alpha(x)]\ .$$

Q.E.D.

Thus if it is exist one element which satisfy (1.1) then all rest minimal elements of Problem 1 must satisfy it.

I illustrate the idea of *α*-functional the next sample.

Let us take some function *f(x)* definite on interval [*a, b*]. Digital values $n \in [a, b]$ are admissible for it. We want find the minimum of this function. The addition member (α –functional) do not change *f(n)* in points *n*, but deforms *f(x)* in gaps between *n* (see fig. 2.1).

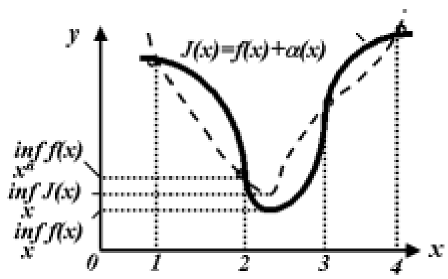


Fig. 2.1.

If *α* – functional is "good", then $\inf_{x \in [a,b]} [f(x) + \alpha(x)] > \inf_{x \in [a,b]} f(x)$. If in addition $\bar{x} = n$, then we get the minimum of Problem 1.

Remark: There are different ways to solve problems by the *α*-functional:

a) You can take the known function as *α*-functional.
b) You can take *α*-functional as unknown function and find it together with the point of minimum.
c) You can take *α*-functional as function *α = α(x,y)* where *α* is known function but *y = y(x)* is unknown function of *x*. You must find it together with the point of minimum.

Let us consider the example. We take as example the non-good the functional which is difficult to solve by conventional method.

Example 1.1. Find the minimum of function

$$I = \frac{4x^2 + 4\pi x + 4.1 + \pi^2}{4(x^2 + \pi x + 1) + \pi^2} \cdot \frac{\sin^5 x - \sin^4 x \cdot \cos x + \sin^2 x \cdot \cos^3 x}{(\sin x - \cos x)(\sin^3 x + \cos^3 x)} \quad in\ X^* = \{x = 0.5\pi n : n = 0, \pm 1, \pm 2, \dots\} \quad (1.2)$$

It is difficult to apply the known methods here because the functional is defined on digital set. The current methods offer only the calculation of all $x \in X^*$. But number of X^* equals infinity and calculation may be meaningless.

Let us to solve this example by the offered method. Take *α(x)* in form

$$\alpha = -\frac{4x^2 + 4\pi x + 4.1 + \pi^2}{4(x^2 + \pi x + 1) + \pi^2} \cdot \frac{0.5 \sin 2x \cdot \cos x}{(\sin x - \cos x)(\sin^3 x + \cos^3 x)}.$$

You can see that *α(x)* = 0 in X^* because for *x* = 0.5π*n* $n = 0, \pm 1, \pm 2, ...,\quad \sin 2x = \sin \pi n = 0$.

Let us to create the general functional

$$J = I + \alpha = -\frac{4x^2 + 4\pi x + 4.1 + \pi^2}{4(x^2 + \pi x + 1) + \pi^2} \cdot \frac{\sin^5 x - \sin^4 x \cdot \cos x + \sin^2 x \cdot \cos^3 x - 0.5 \sin 2x \cdot \cos x}{(\sin x - \cos x)(\sin^3 x + \cos^3 x)}.$$

Here the variable *x* is uninterrupted and - ∞ < *x* < ∞ (set *X*)

The additive *α(x)* allows to change the functional (1.2) to simple form

$$J = \frac{4x^2 + 4\pi x + 4.1 + \pi^2}{4(x^2 + \pi x + 1) + \pi^2} \cdot \frac{(\sin^2 x - \cos^2 x)(1 - \sin x \cdot \cos x)\sin x}{(\sin^2 x - \cos^2 x)(\sin^2 x - \sin x \cdot \cos x + \cos^2 x)} = \left(\frac{0.1}{4 + (2x + \pi)^2} + 1 \right) \sin x .$$

This general functional is simple. His minimum may be found the conventional method of theory the function one variable. Here $\bar{x}$ = - π/2 , $\bar{x} \in X^*$ for $\bar{n} = 1, \quad \bar{I} = -1.025$. Consequently, that is absolute minimum (and sole) of initial functional (1.2).

We can apply an analogical method for finding of minimum on *x* the next functional

$$I = \cos^2 \varphi + 0.5 \cos 2x \cos 2\varphi - 2 \cos x \cos \varphi \cos(x + \varphi) + 0.5 - 0.1 e^{-x^2}, \quad X^* = \{x = 0.5\pi n : n = 0, \pm 1, \pm 2, ...\}.$$

Here *φ* is given, *x* is digital. Let us take $\alpha = -0.5 \sin 2x \sin 2\varphi$. After this we can change our functional *J* = *I* + *α* to simple form: $J = -0.1 e^{-x^2} + \sin^2 x$. The point of absolute minimum this task (Problem 2) is $\bar{x}$ = 0 . This point is in allowable set X^* for $\bar{n} = 0$. That means $\bar{n} = 0$ is point of the absolute minimum od the initial Problem 1.

The reader can think: if the allowable numerical set is limited we can use the conventional Lagrange's method [7]. Let us show: that is not correct.

Example 1.2. Find minimum of functional:

$$I = x^3 - 3x^2 + x \quad on \quad X^* = \{x = 0, x = 3\}. \qquad (1.3)$$

Let us to write the Lagrange's function

$$F = x^3 - 3x^2 + 2x + \lambda_1 x + \lambda_2 (x - 3),$$

where λ_1 , λ_2 are LaGrange's factors. Find the first derivative

$$F' = 3x^2 - 6x + 2 + \lambda_1 + \lambda_2 .$$

Substitute to here *x* = 0 , *x* = 3 and write the equations $F'(0) = 0, \quad F'(3) = 0.$ We find from these equations λ_1 , λ_2 . Find the second deviation $F'' = 6x - 6.$ When *x* = 0 the function $F''(0) = -6 < 0.$ When *x* = 3 the function $F''(3) = 12 > 0.$ Consequently *x* = 0 is the point of maximum, *x* = 3 is the point of minimum. Let us check up. Substitute *x* = 0 and *x* = 3 in (1.3). We find *I*(0) = 0, *I*(3) = 6 .

We see the LaGrange's method gives the opposed result: it declare the point of minimum as the point of maximum, but the point of maximum as the point of minimum. In here it is violate one condition of LaGrange's method: The number of additional equations is more of number of variables. This example is shows: this violation for LaGrange's method is unacceptable.

Let us to solve this example by the offered method. Take the *α(x)* in form

$\alpha = x(x\text{-}3)(2/3\text{-}x)$.

Then

$$J = I + \alpha = x^3 - 3x^2 + 2x + x(x-3)(2/3 - x), \quad J' = 4/3x = 0, \quad \bar{x} = 0 \in X^*, J'' = 4/3 > 0 .$$

From Consequence 1 the point $\bar{x} = 0$ is absolute minimum of functional (1.3). That shows the method of *α* – functional has more application then the the LaGrande's method.

Example 1.3. Find minimum of integral

$$I = \int_{-10^{-3}}^{a} (\ln \mathrm{tg}\, t - 10^{-3}) dt \quad on \quad X^* = \{a = 10^{-3} \pi n: \ n = 1, 2, ..., 400\} \qquad (1.4)$$

Here the interval of integration is discrete. The direct search is difficult because integral (1.4) cannot be presented by simple function and it not have of tabulations.

Let us to find *α*-functional in form: $\alpha = -10^{-6} \sin 10^3 a$. You see on X^* the function *α*(*x*) = 0. Further

$$J = I + \alpha = \int_{10^{-3}}^{a} (\ln tg\ t - 10^{-3}) dt - 10^{-6} \sin 10^3 a\ ,$$

$$J_a^{'} = \ln tg\ a - 10^{-3} - 10^{-3} \cos 10^3 a = 0\ , \quad \bar{x} = \pi/4 \in X^* \quad for \quad \bar{n} = 250\ , \qquad (1.5)$$

$$J'' = \frac{2}{\sin 2a} + \sin 10^3 a \quad .$$

As $10^{-3} < x < 0.4\pi$, then $J'' > 0$ into this interval. That means the root is single and $\bar{n} = 250$ is point of the absolute minimum.

Analogically we find the minimum of other integral which cannot be presented in simple functions

$$I = -\int_0^a [\sin(t^3) + 10^{-5}\sqrt{\pi}] dt \quad on \quad X^* = \{a = 10^{-3}\sqrt{\pi} n : n = 0, 1, ..., 1.5 \cdot 10^3\} \, . (1.6)$$

Here is $\alpha = 10^{-3} \sin 10^{-8} \sin 10^3 \sqrt{\pi} a\,; \quad \bar{n} = 1000.$

Example 1.4 . Find the minimum of integral

$$I = \int_{\pi/2}^{\pi} \left(\frac{\cos at}{t} + 20a^3 \right) dt \quad on \quad X^* = \{a = 10^{-3} n : n = 0, \pm 1, \pm 2, ...\}\, . \qquad (1.7)$$

Here the under integral function is discrete. The integral from this function cannot be presented as elementary functions.

Let us take $\alpha = 10^{-3} \sin^2 10^3 \pi a\,, \quad J = I + \alpha$. Then

$$J_a^{'} = I_a^{'} + \alpha_a^{'} = \int_{\pi/2}^{\pi} (-\sin at + 40a) dt + 2 \cdot 10^{-4} \pi \sin 2 \cdot 10^3 \pi a =$$
$$= -\frac{2}{a} \sin \frac{3}{4} \pi a \cdot \sin \frac{\pi}{4} a + 20\pi a + 10^{-4} \pi \sin 2 \cdot 10^3 \pi a \qquad (1.8)$$

This derivative not exist for $\overline{a} = 0 \in X^*$.

$For \quad a \geq 0, \quad J' > 0; \quad for \quad a < 0, \quad J' < 0; \quad (or \quad J'' > 0 \quad for \quad \forall a \neq 0)$.

Consequently $\overline{n} = 0$ is point of absolute minimum.

B) Consider the case when the point of optimum $x^* \in X^*$ not exist, but exist the sequence such that $\lim_{n\to\infty} I(x_n) = m$. This sequence is named the minimizing sequence (see §5 of Ch.1).

Similarly point A we can show that consequence 1 can be generalized in this case.

Consequence 1'. Let us *α(x)* = 0 only on *X**, For minimizing sequence $\{x_n\} \subset X^*$ is necessary and sufficient the existing of function *α(x)* such that

$$\lim_{n\to\infty}[I(\alpha_n) + \alpha(x_n)] = \inf[I(x) + \alpha(x)], \quad x \in X. \tag{1.9}$$

The sufficiency of this consequence is same the lemma in [2] and *J(x)* = *L* in [2].

We can generalize remark 3 of item 1 in this case: If exist *α* function and one sequence $\{x_n\} \subset X^*$ which satisfy (1.9), then the any sequence $\{x_n\} \subset X^*$ which satisfy (1.9) is the minimizing sequence. And on the contrary any the minimizing seuence satisfy the condition (1.9).

2. α – function in Banach space.

Let us to apply Theorem 1.2 to optimal problem is described in Banach space by equation

$$\frac{dx}{dt} = f(x,u), \quad t_1 \leq t \leq t_2, \quad x(t_1) = x_1, \quad x(t_2) = x_2, \tag{1.10}$$

where *x, f(x,u)* – element complete linear normed space X_1 and X_2 respectively and $X_2 \subset X_1, \quad t \in = [t_1, t_2] = T$ is segment of real axis.

Let us name the permissible control the measurable limited function (in term [1], p.85)with value $u \in U$, where *U* is set in arbitrary topological space. In particular the set *U* may be metric, closed and limited. Let us assume that for any control *u(t)*the equation (1.10) has single solution *x(t)* with $x(t) \in X_1$ for almost all $t \in [t_1, t_2]$, where *x(t)* is continuous almost everywhere differentiable on function on $t \in [t_1, t_2]$.

Operator *f(x,u)* is defined on the direct product *X×U* . One is continuous and bounded. Boundary conditions are given t_1, t_2, $x(t_1) = x_1$, $x(t_2) = x_2$.

State the problem: Find the admissible control which transfers the system from given initial state in given final state with function

$$I = \int_{t_1}^{t_2} f_0(x, y)dt \qquad \textbf{(1.11)}$$

has a minimum.

Let us the set of the measurable functions *u (t)* is denoted *V*: set of the continuous, almost everywhere differentiable on (t_1, t_2) the functions *x (t)* is denoted *D*. Set of couple x(t), u(t) having named over properties and almost all satisfied the equation (1.10), we name *admissible* and denote Q. It is obvious $Q \subset D \times V$.

Assume $\psi = \psi(t, x)$ **is the some unequivocal continuous differential** function defined on *X×T* . We name it the *characteristic function* . We will find the α – function in form

$$\alpha = \int_{t_1}^{t_2} \psi_x * [\dot{x} - f(x,u)]dx \qquad \textbf{(1.12)}$$

Here $\psi_x = \dfrac{\partial \psi}{\partial x}$ **is particular deviation of Freshe. One is linear function. The * is sign of composition. Obvious that request of α-function is performed.**

Compose the generalized function *I = J + α* and produce the function $\dot{\psi} = \psi_x \dot{x} + \psi_t$ **we get**

$$J = \psi[t_2, x(t_2)] - \psi[t_1, x(t_1)] + \int_{t_1}^{t_2} (f_0 - \psi_t - \psi_x \circ f)dt = \psi_2 - \psi_1 + \int Bdt \quad , \qquad \textbf{(1.13)}$$

where $B = f_0 - \psi_t - \psi_x \circ f$. Because the set *Q* is different from the set *D×V* only that couple *x(t), u(t)* satisfy almost every where (1.10). For α-function in form (1.12) with according of Theorem 1.2 we can the initial Problem 1 (find the minimum (1.11) on *Q*) replace the Problem 2 – find minimum (1.13) on the broader set *D×V*. In this set the *x(t), u(t)* not bind the equation (1.10). So we have

$$\bar{J} = \psi_2 - \psi_1 + \inf_{x(t) \in D, u(t) \in V} \int_{t_1}^{t_2} B(t, x, u)dt \, . \qquad \textbf{(1.14)}$$

Theorem 1.6. If function $\bar{u}(t)$ **getting from solution of problem** $\inf\limits_{x(t) \in D, u(t) \in V} \int_{t_1}^{t_2} Bdt$ **is** $\bar{u}(t) \in V$ **, that it is same almost everywhere the function getting from solution the problem** $\inf\limits_{x(t) \in D, u(t) \in V} \int_{t_1}^{t_2} Bdt$ **and**

$$\inf_{x(t) \in D, u(t) \in V} \int_{t_1}^{t_2} Bdt = \inf_{x(t) \in D} \int_{t_1}^{t_2} \inf_{u \in V} Bdt \qquad \textbf{(1.15)}$$

Proof. **Assume the contrary:** $B(u^*) \neq \inf\limits_{u \in U} B(u)$ **on subset of interval [t_1, t_2] with measure not equal zero. In this case** $B(u^*) > B(\bar{u})$ i.e. $\int_{t_1}^{t_2} B(u^*)dt > \int_{t_1}^{t_2} B(\bar{u})dt$ **on the subset. This contrad**ict: the function $u*(t)$ **made the minimum for integral** $\int_{t_1}^{t_2} Bdt,$

From requirement (1.14) and Theorem 1.6 we have

$$\bar{J} = \psi_2 - \psi_1 + \inf_{x(t)\in D} \int_{t_1}^{t_2} \inf_{u\in V} B(t,x,u)dt \qquad \textbf{(1.16)}$$

If function $\alpha[x(t),u(t)]$ **is such that absolute minimum of Problem (1.16):** $\vec{x}(t),\ \bar{u}(t) \in Q$ **, then ααaccording to Theorem 1.1 functions** $\bar{x}(t),\bar{u}(t)$ **are absolute minimum of the initial Problem.**

So, we proofed

Theorem 1.7. To couple function were the absolute minimum the function *I* , it is sufficient the existing the characteristic function $\psi(t,x)$ **such that**

1) $B(t,x,\bar{u}) = \inf_{u\in U} B(t,x,u)$; 2) $\int_{t_1}^{t_2} B(t,x,\bar{u})dt = \inf_{x(t)\in D} \int_{t_1}^{t_2} B(t,x,\bar{u})dt$; 3) $\vec{x}(t),\ \bar{u}(t) \in Q$ **; (1.17)**

In particular, if take $\psi = p(t) \circ h$ **, where p(t) is linear function** $h \in X_1$**, then from item 1 and stationary condition item 2 [1.17] we get**

$$H(t,x,\bar{u}) = \sup \sup_{u\in U} H(t,x,u)\ , \quad \dot{p}(x) = -\frac{\partial H}{\partial x}\ , \qquad \textbf{(1.18)}$$

where $H = p(t) \circ f(x,u) - f_0(x,u)$. **Assumed** $\partial H / \partial x$ **is Frechet derivative, which is continuous.As we see the necessary condition of Problem 2 following from (1.17) is same the necessary condition of Pontriagin principe of maximum generalized in Banach spaces.**

3. Design of α-function for allowable subset of two function connected by logical conditions

Assume two functions *F₁(x*) and *F₂(x)* are refinished on the set *X*. Allowable are only points $x \in X$ and functions F_1 and F_2 which are connected the logical conditions. Assume $F_1(x) = 0$ is "true" and $F_2(x) \neq 0$ is "false'. The five main logical connections (⟷, y, v, ∧, ~) (⟷, y, ∨, ∧, ~) are presented in next tables:

F_1	F_2	$F_1 \leftrightarrow F_2$
t	t	t
t	f	f
f	t	f
f	f	t

Double implication

F_1	F_2	$F_1 y F_2$
t	t	f
t	f	t
f	t	t
f	f	f

disjunction in the exclusive sense

F_1	F_2	$F_1 \vee F_2$
t	t	t
t	f	t
f	t	t
f	f	f

disjunction in the sense of a non-exclusive

F_1	F_2	$F_1 \wedge F_2$
t	t	t
t	f	f
f	t	f
f	f	t

Conjunction

F	*p*
t	*f*
f	*t*

Denial

We will use the symbol:

$$\text{sign}\,F = 1 \quad if \quad F > 0,$$
$$\text{sign}\,F = 0 \quad if \quad F = 0,$$
$$\text{sign}\,F = -1 \quad if \quad F < 0,$$

In this case the α-function we can search in form:

$$1)\ X^* = \{x : F_1(x) \leftrightarrow F_2(x)\}, \quad \alpha = (p_1F_1 + p_2F_2)[1 - |\,\text{sign}(F_1F_2)\,|],$$
$$2)\ X^* = \{x : F_1(x)\ y\ F_2(x)\}, \quad \alpha = p_1F_1F_2 + p_2[1 - |\,\text{sign}(F_1^2 + F_2^2)\,|],$$
$$3)\ X^* = \{x : F_1(x) \vee F_2(x)\}, \quad \alpha = pF_1F_2\ ,$$
$$4)\ X^* = \{x : F_1(x) \wedge F_2(x)\}, \quad \alpha = p_1F_1 + p_2F_2\ ,$$
$$5)\ X^* = \{x : F_1(x) \sim F_2(x)\}, \quad \alpha = (p[1 - |\,\text{sign}\,F\,|],$$

Here *p*, p_1, p_2 are some function *x* .

It is using these five connections we can create all other complex logic statements.

§2. The general principle of reciprocity the optimization problems

Let us suppose we want to solve the optimal problem Ch.1 §4 p.4 :

$$I = f_0(x)\,, \quad f_i(x) = 0, \quad i = 1,2,\dots,m\,, \qquad (2.1)$$

Design general function in form

$$J = \sum_{i=0}^{i=n} \lambda_i(x,y)\, f_{i,}(x)\ , \qquad (2.2)$$

where $\lambda_i(x,y)$ arbitrary functions of *x, y*.

Assume $\bar{x}(y)$ is absolute minimum (2.2) on *X*.

<u>The general principle of reciprocity the optimization problems</u>

1. For any $y \in Y$ the point of an absolute minimum of the function *J* (2.2) is the point of the absolute minimum any function

$$\lambda_j(x.y)\, f_j(x)\,, \quad j = 0.1,\dots,m \quad (\text{no sum for } j) \qquad , \qquad (2.3)$$

for limits in form

$$\lambda_i(x,y) = \lambda_i(\bar{x}(y), y)\, f_i(\bar{x}(y))\,, \quad i = 0,1,\dots,m\,, \quad i \neq j\,, \quad (\text{no sum for } i)\,. \qquad (2.4)$$

Any numbers of equality (2.4) you can change by non-equalities

$$\lambda_i(x,y) \le \lambda_i(\bar{x}(y),y)\, f_i(\bar{x}(y)\,)\,. \qquad (2.5)$$

2. For any $y \in Y$ the point of the absolute minimum of the function *J* (2.2) is point of the absolute minimum any sum the functions

$$\sum_j \lambda_j(x,y)\, f_j(x) \qquad (2.3)'$$

for restrictions absent in sum (2.3)

$$\lambda_i(x,y) = \lambda_i(\bar{x}(y),y)\, f_i(\bar{x}(y)\,, \quad i = 0,1,...,m\,, \quad i \ne j\,, \quad (\text{no sum for } i)\,. \qquad (2.4)'$$

Any numbers of equality (2.4)' you can change by non-equalities (2.5).

Proof.
1) For any function (2.3) for conditions (2.4) the Theorem 1.2 is made. The point $\bar{x}(y)$ is point of its absolute minimum. As every function reaches the global minimum, obvious, the change equality (2.4) by restrictions (2.5) not influence to minimum. The point 2 is proofed similarly. Principle is proved.

Consequence 1.

Magnitude $J(\bar{x}(y), y)$ is the lower estimation of any function from (2.3), (2.3)' if part or all equalities (2.4), (2.4)' change equalities in form

$$\lambda_i(x,y)\, f_i(x) = 0 \qquad (2.6)$$

Consequence 2. In case corresponded (2.6) the absolute minimum of any functions (2.3) are located in set

$$M_j(y) = \{x : \sum_{\substack{i=1\\ i\ne j}}^{m} \lambda_i(x,y)\, f_i(x) \ge \sum_{\substack{i=1\\ i\ne j}}^{m} \lambda_i((\bar{x}(y),y)\, f_i(\bar{x}(y)) \qquad (2.7)$$

Consequence 3. If possible the solution of Problem (2.1) by Algorithm 4, there are *y* such that

$$\lambda_i((\bar{x}(y),y)\, f_i(\bar{x}(y) \le 0 \quad (\text{no sum for } i) \qquad (2.8)$$

From the existence of solutions (2.1) follows that $f_i(x) = 0$. So $\bar{\lambda}_i \bar{f}_i$ is minimum, than (2.8) is obvious.

§3. Applications α-function to well-known Problems of optimization

1. Problem the searching of conditional extreme the function of the limited number variables.
It is given

$$I = f_0(x)\,, \quad f_i(x) = 0\,, \quad i = 1,2,...,m < n \qquad (3.1)$$

Here x is n-dimensional vector given in some numerical open region of n-dimensional space X^*.

Let us take the α-function in form

$$\alpha = p_i(x)\, f_i(x)\,, \quad i = 1,2,...,m \qquad (3.2)$$

(repeated indexes means summarization). Here $p_i(x)$ are functions x, given on X:

$$X^* = \{x : \textstyle\sum_{i=1}^{m} |\, f_i(x) | = 0\},\ X^* = X.$$

Let us to design generalized functional $J(x) = f_0(x) + \alpha(x)$ take some $p_i(x)$ and sole the problem $\inf J(x)\,, \quad x \in X$. From this solution the Problem 2, according Theorems §1, we can get the following information about Problem 1:

1) If $\bar{x} \in X^*$, than $\bar{x}$ is absolute minimum of Problem 1 (consequence 1, §1).

2) If $\bar{x} \notin X^*$, then:

a) $J(\bar{x})$ is the lower estimation of function $f_o(x)$ on X^* (Theorem 1.4).

b) For $\alpha(\bar{x}) > 0$ x^* is located in set $P = \{x : \alpha(x) \le \alpha(\bar{x})\}$ (consequence 3, §1).

c) For $\alpha(\bar{x}) < 0$ x^* is located in set $M = \{x : \alpha(x) \ge \alpha(\bar{x})\}$ (consequence 4, §1).

d) Set $N^* = N \cap X^*$ where $N = \{x : 2f_0 + \alpha \le 2\bar{f}_0 + \overline{\alpha}\}$ contains the equal or worse solutions (Theorem 1.5).

As we see even if $\bar{x} \notin X^*$ our computation is useful. We received the lower estimation and narrow the region for searching of the optimal solution. Take row of α_i we can get the solution one of the Problems ***a, b, c, d*** or facilitate the solution of Problem ***a*** (see Ch, 1, §1).

Look your attention: the offered method does not require continuity and differentiability of the functions $f_0(x)$, $f_i(x)$ in contrast to the classical method of Lagrange multipliers. The method can be applied to non analitical function, for example, to the functions definished on the discret set and extremal problems of the combinatorics (see Ch. 10).

2. Application the Theorems §1 to optimal problems described the conventional differential equations.

Assume the moving of object is described by system of the differential equations

$$\dot{x}_i = f_i(t,x,u)\,, \quad i = 1,2,...,n\,, \quad t \in T = [t_1, t_2]\,, \qquad (3.3)$$

where *x(t)* –*n*-dimensional continuous piecewise differentiable function, $x \in G(t)$; *u(t)* – *r*-dimensional functions continuous everywhere on *T*, except limited number of points where one can have discontinuity of the first kind $u \in U(t)$. Boundary values t_1, t_2 are given, $x(t_1), x(t_2) \in R$.

Optimal function is

$$I = F(x_1, x_2) + \int_{t_1}^{t_2} f_0(t, x, u)dt, \quad x_1 = x(t_1), \quad x_2 = x(t_2)\,. \tag{3.4}$$

Functions $F(x_1,x_2)$, $f_i(x,u,t)$, $i = 0,1,\dots,n$ are continuous, $F(x_1,x_2) > -\infty$. Set of the continuous almost everywhere differentiable functions *x(t)* with $x \in G(t)$ we designate *D*. Set of the piecewise continuous (they can have the discontinuity of the first kind) functions u(t) such that $u \in U(t)$ we designate *V*. Couple *x(t), u(t)* have named over properties and almost everywhere satisfy the equations (3.3) we name allowable and designate *Q*, $Q \subset D \times V$.

Enter in our research *n* single-valued functions $\lambda_i(t.x)$ $i = 1,2,\dots,n$. which are continuous and have continuous derivatives on T×G. Let us to take the α-function in form

$$\alpha = \int_{t_1}^{t_2} \lambda_i(t,x)[\dot{x} - f_i(t,x,u)]dt \tag{3.5}$$

It is obvious α = 0 on *Q*. Let us design the general function *J = I + α* , integer the term $\lambda_i \dot{x}_i$ by part and exclude $\dot{x}_i$ by (3.3). We get

$$J = F + \lambda_i x_i \big|_{t_1}^{t_2} + \int_{t_1}^{t_2} [\, f_0 - (x_j \frac{\partial \lambda_j}{\partial x_i} + \lambda_i)\, f_i - x_i \frac{\partial \lambda_i}{\partial t}]dt \tag{3.6}$$

Designate

$$a = F + \lambda_i x_i \big|_{t_1}^{t_2}, \quad B = f_0 - (x_j \frac{\partial \lambda_j}{\partial x_i} + \lambda_i)\, f_i - x_i \frac{\partial \lambda_i}{\partial t}$$

Apply to (3.6) Consequence 1 §1. Here the *Q* is *X** and *D*x*V* is *X* (see Consequence 1 §1). Since now the couple of functions *x(t), u(t)* from *D*x*V* (having ends in *R* for condition $\bar{x}(t) \in D, \bar{u}(t) \in V,\, x_1 = x(t_1),\, x_2 = x(t_2)$) are not connected by the equations (3.3) we can write

$$\inf_{D \times V} (A + \int_{t_1}^{t_2} Bdt) = \inf_{x_1, x_2 \in R} A + \int_{t_1}^{t_2} \inf_{x \in G, u \in U} Bdt$$

and final

$$\bar{J} = \inf_{x_1, x_2 \in R} A + \int_{t_1}^{t_2} \inf_{x \in G, u \in U} Bdt \tag{3.7}$$

So we proofed the **Theorem 3.**1:
The couple vector-function $\bar{x}(t), \bar{u}(t)$ will be point of absolute minimum of function (3.4) if it is exist *n* differentiable *λ*$\lambda_i(t,x)$ such that:

$$1)\ \bar{B} = \inf_{x \in G, u \in U} B, \quad 2)\ \bar{A} = \inf_{x_1, x_2 \in R} A > -\infty, \quad 3)\ \bar{x}, \bar{u} \in Q \qquad (3.8)$$

Note: That is sufficient condition only. That cannot be a necessary condition because we don't know advance about an existence of $\lambda(t,x)$.

From (3.8) it is follow: if we find at least one solution of an equation in particular derivations having *n*-unknown functions $\lambda_i(t,x)$:

$$\inf_{u \in U} [f_0 - (x_j \frac{\partial \lambda_j}{\partial x_i} + \lambda_i) - x_i \frac{\partial \lambda_i}{\partial t}] = 0, \qquad (3.9)$$

for boundary condition *A* = const, than points 1, 2 of the Theorem 3.1 will be executed. Any unsuccessful $\lambda_i(t,x)$ (if $\bar{x}(t), \bar{u}(t) \notin Q$) with according Theorem 1.4 gives the lower estimation of the global minimum.

Assume, for example, $x_n \neq 0$*. Substitute them in (3.7), we get the result published in work [2]**, (condition Bellman-Pikone):

$$\bar{J} = \inf_{x_1 \in G_1,\ x_2 \in G_2} \Phi - \int_{x_1}^{x_2} \sup_{x \in G, u \in U} R(t,x,u) dt, \qquad (3.10)$$

Here $\Phi = F + \varphi_{t_1}^{t_2}$, $R = \varphi_t + \varphi_{x_i} f_i - f_0 = -B.$

----

* This limitation is not important becouse any $x_i \neq 0$ in [t_1,t_2].

** Note: in given method (in difference from [2]) not request a priory assamption about existing the single potensial function *φ(t,x)* such that $\varphi_{xi} = \lambda_i$.

Sometimes it is more comfortable take function *φ(t,x)*
or in other terms (see [4]) *ψ(t,x).* Then *A, B* are written:

$$A = F + \psi_2 - \psi_1, \quad B = f_0 - \psi_{x_i} f_i - \psi_t , \qquad (3.11)$$

And Theorema 3.1 is same with [2], (see also [3]).

Function α for given task wecan define also the next way. Take some function *ψ(t,x)*. Than

$$\alpha = \int_{t_1}^{t_2} \psi_{x_i} [\dot{x} - f_i(t,x,u)] dt$$

Integrate but parts the first member we get

$$\alpha = \psi |_1^2 - \int_{t_1}^{t_2} (\psi_{x_i} f_i + \psi_t) dt$$

Note: 1. Theorema 3.1 is corrected and in notations (3.8) p.1:

$$\int_{t_1}^{t_2} B dt = \inf_{x(t) \in B} \int_{t_1}^{t_2} \inf_{u \in U} B dt.$$

This form is offered in [4]. Difference between these forms is important in consideration the second variation, conditions in angle points and in some other cases. Let us takethe last corrected form of V. Krotov optimization [8] (problem of speed):

Example 3.1. Find minimum t_2 in task:

$$I = \int_{t_1}^{t_2} dt, \quad \dot{x} = u, \quad |u| = 1, \quad x(0) = 1, \quad x(t_2) = 0.$$

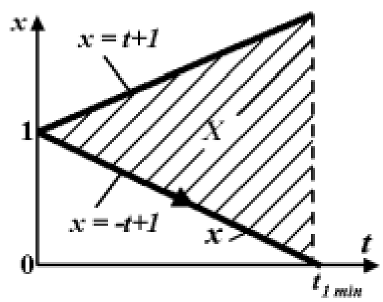


Fig.2.2.

If we take φ = 0, we get R = -1. Consequently $\sup_{x,u} R$ is reached in ANY curve, for example, u = - 0.01 (I = 100). In case whem *min* forward integral for ψ = 0 we have

$$\inf_{x(t) \in D} \int_{t_1}^{t_2} dt = \inf_{x(t) \in B} t_2[x(t)].$$

Since the set all serves with bounded derivative $|\dot{x}| \le 1$ for $x(0) = 1$ located between lines $x = t - 1$, $x = -t + 1$ (Fig. 2.2), we get $\bar{x} = 1 - t, \quad \bar{u} = -1 \quad and \quad I = t_{1,\min} = 1.$

Notes: 1. As set B we can take a set $\{x(t)\}$ with bounded derivative $\dot{x}_i \in X_i = \{f_i(t, x, u) : u \in U\}$. This narrowing can help in finding of optimal solution.

2. Note 3 §1 in given case has the following view: If exist the function *ψ(t,x)* and at list one allowable couple $\bar{x}(t), \bar{u}(t)$, satisfying (3.8). That any other couple satisfying (3.8) is minimum of problem 1 and any allowable minimum the problem 1 satisfy p.1, 2 (3.8).

3. If t1, t2 are not fixed, we can show that point 1, 2 (3.8) are:

$$1)\ \overline{B} = \inf_{x\in G,\, u\in U} B = 0, \quad 2)\ \overline{A} = \inf_{t_1,t_2,x_1,x_2\in R} A > -\infty$$

We can satisfy the condition *inf B* = 0, if we take $\psi = \varphi(t,x) + y_{n+1}$ and

$$\dot{y}_{n+1} = f_0 - \varphi_{x_i} f_i - \varphi_t \cdot$$

4) Theorem 3.1 is particular case of more common theorem 2.1 considered in Chapter Ш.

Assume we take some $\lambda_i(t,x)$ (or $\psi(t,x)$).

Theorem 3.2. Assume *F* = 0 and solved the problem $\inf_{x,u} B$. That:

1) Set $N = \{t,x,u : B + f_0 \le \overline{B} + \bar{f}_0,\ t\in T\}$ contains same and better solutions of Problem 1;
2) Set $P = \{t,x,u : B - f_0 \le \overline{B} - \bar{f}_0,\ t\in T\}$ contains same and worse solutions of Problem 1.

Proof: 1) Deduct $B \ge \overline{B}$ from inequality $B + f_0 \le \overline{B} + \bar{f}_0$. We get

$$f_0 \le \bar{f}_0 \quad \text{on} \quad T, \quad \text{i.e.} \quad \int_T f_0 dt \le \int_T \bar{f}_0 dt \cdot$$

2) Deduct $B \ge \overline{B}$ from inequality $B - f_0 \le \overline{B} - \bar{f}_0$. We get

$-f_0 \le -\bar{f}_0 \quad \text{on} \quad T, \quad \text{i.e.} \quad \int_T f_0 dt \ge \int_T \bar{f}_0 dt \cdot$ Theorem is prooved (QED).

Let us take instead function (3.4) simpler function $\int_T B_1(t,x,u)dt$ (here B1 is given function). Than

Theorem 3.3. Assume *F* = 0 and solved the problem $\bar{J}_1 = \inf \int_T B_1(t,x,u)dt \text{ on } Q \cdot$ Than:

3) Set $N = \{t,x,u : B_1 + f_0 \le \overline{B}_1 + \bar{f}_0,\ t\in T\}$ contains the same and better solutions of Problem 1;
4) Set $P = \{t,x,u : B_1 - f_0 \le \overline{B}_1 - \bar{f}_0,\ t\in T\}$ contains the same and worse solutions of Problem 1.

Proof: 1) From *N* we have the inequality $\int_T (f_0 + B_1)dt \le \int_T (\overline{B}_1 + \bar{f}_0)dt$. Deduct from this inequality the inequality $\int_T B_1 dt \ge \int_T \overline{B}_1 dt$. We get $\int_T f_0 dt \le \int_T \bar{f}_0 dt$.

2) From *P* we have the inequality $\int_T (B_1 - f_0)dt \le \int (\overline{B}_! - \bar{f}_0)dt$. Deduct $\int_T B_1 dt \ge \int \overline{B}_1 dt$ from this inequality. We get $\int_T f_0 dt \ge \int_T \bar{f}_0 dt$. Theorem is prooved (QED).

Consequence. If set *P* cover the set $T\times G\times U$ (or reachability set) and $\bar{x},\bar{u} \in Q, \quad \text{then} \quad \bar{x},\bar{u}$ are absolute minimum of Problem 1.

Note. Delete part equation (3.1) or (3.2) [in case (3.2) x_i correspoded deleted equations became the control in the rest equations]. Then getten solution is the low estimation of initial Problem as it is follow from principle of expension [5]: $I(x) \geq I(\bar{x})$ and $I(x,u) \geq I(\bar{x},\bar{u})$, where $\bar{x}(t), \bar{u}(t)$ are absolute minimum "truncated" task.

When right parts of equations (3.3), (3.4) do not depent clearly from x(t), we can stand out not only set *N,P* but the set *M*. It is correct the following theorema

Theorem 3.4. *Assume* $F \geq 0$, *ends x(t) is free, the right parts of equations (3.3), (3.4) depent only from t, u , i.e.:* $f_i = f_i(t,u)$ *i = 0,1,…,n. and solved task* $\inf_{x,u} B_1(t,u)$. *Than:*

1) Set $M = \{t,u : B_1 - f_0 \geq \bar{B}_1 - \bar{f}_0, t \in T\}$ *contains the absolute minimum of Problem 1;*

2) Set $N = \{t,u : B_1 - f_0 \geq \bar{B}_1 - \bar{f}_0, t \in T\}$ *contains the same and better solutions of Problem 1;*

3) Set $P = \{t,x,u : B_1 - f_0 \leq \bar{B}_1 - \bar{f}_0, t \in T\}$ *contains the same and worse solutions of Problem1*.

Proof for sets *N*, *P* full equally with the proof of Theorem 3.2. Proof for *M* follows from discontinuity *u(t)* and depends the right parts of equation only from *u*.

3. Task the dynamic programming of Bellman

Assume there is physical system *S*. The control of this system separated in *m* steps. On every *i* step we have the control U_i. Using this control we transver our system from allowable stand S_{i-1} getted in (*I* - 1) step in new allowable stand $S_i = S_i(S_{i-1}, U_i)$. This transwer is bounded by some conditions. The purpose is minimum function

$$W = \sum_{k=1}^{n} w_k$$

Let us to biuld the common function

$$J_i = W_i + \alpha, \text{ where } W_i = \sum_{k=1}^{n} w_k, \; i = 1,2,\ldots,m \, .$$

In this case we can change the task of the conditional minimum inf Wi in the task of direct minimum $\inf_V J_i$. If the limitations are absent or they allow the select U_k in every step to make with associated conditions, then from $\alpha = 0$ in the admisseble elements we get the Bellman equation [6].

$$\overline{W_i}(S_{i-1}) = \min_{U_i}\{W_i(S_{i-1}, U_i\}, \quad i = 1,2,\ldots,m.$$

3. Application α-function for solution the problems with distributed parameters

Let us consider about absolute minimum the Problem with distributed parameneers

$$I(x,u)=\int_P f_0(t,x,u)dt+F(x(\tau))' \qquad (3.12)$$

where $t=(t_1,t_2,...,t_m), \quad x=(x_1,x_2,...,x_n), \quad u=(u_1,u_2,...,u_r)$ are elements of vector space *T, X, U* * respectively. *P* is closed area in space *T*, bounded continuous piecewise smooth, fixed hypersurface *S*. On *S* the *t = τ*. *P** is internal part this area, functions $x_i(t)$ on *P* are absolute-continious, $u_\alpha(t)$ are measurable on *P* and have velues from area *U*, which can be closed and bounded.

Functions *x(t), u(t)* satisfy almost everywhere the system $n \cdot m$ independed differensial equetions with particulal deviations

$$\frac{\partial x_i}{\partial t_j}=f_j^i(t,x,u), \quad i=1,2,...,n; \quad j=1,2,...,m \qquad (3.13)$$

Funsions $f_j^i,\ f_0$ are continiouly together with its particular derivitives the first order. The function *x(t), u(t)* we name allowable if they satisfy the named above conditions (set *Q*).

<u>Statement og Problem</u>: Find couple function *u(t), x(t)*, which give the function *I* (3.12) the minimal value.

Add to system (3.13) the integrability condision:

$$\varphi^\gamma=\frac{\partial f_j^i}{\partial t_k}-\frac{\partial f_{ki}^i}{\partial t_j}=0, \quad i=1,2,...n; \quad j,k=1,2,...,m; \quad k>j. \qquad (3.14)$$

Not difficult to calculate, that number of difficalt equation (3.14) may be $0.5(m-1)mn, \quad i.e. \quad \gamma=1,2,...,0.5(m-1)mn$ (number of combinations $C_m^2 n)$. For simplicity we will assume: all functions φ^γ in (3.14) contain *u* and these *u* may be find from (3.14) Assume the number of independed equations (3.14) are less *r*.

Let us lead to consider *m*-dimentional function $\psi(t,x)=\{\psi^1,\psi^2,...,\psi^m\}$. The components of this function $\psi^j(t,x)$ $j=1,2,...,m$ are continious and have the continuous partial derivatives almost everywhere in *T*.

Name this function – charasteric function. Let us lead also the integrable vector-function

$\lambda_1(t), \lambda_2(t),...,\lambda_p(t)$.

Let us take α- function in form

$$\alpha=\int_S \psi^j(\tau,x)\cos(n,t^j)d\tau-\int_P(\psi_{t_j}^j+\psi_{x_i}^j f_j^i+\lambda_\gamma\varphi^\gamma)dt' \qquad (3.15)$$

Where *n* is outer normal to surface *S, dτ* is element surface *S*. We present the function J = I + α in form

$$J=A+\int_P Bdt, \quad where \quad A=\int_S \psi^j(\tau,x)\cos(n,t^j)d\tau, \quad B=f_0-\psi_{t_i}^j+\psi_{x_i}^j f_j^i+\lambda_\gamma\varphi^\gamma. \qquad (3.16)$$

Theorem 3.5. *Assume* $u(t) \in V$. *In order to couple u(t), x(t) will be the absolute minimum the purpose function (3.12) it is sufficiently* exicting of α-fuction (3.15) such that*

$$1)\ \bar{B} = \inf_{x,u \in U} B(t,x,u), \quad 2)\ \bar{A} = \inf_{x(\tau)} B > -\infty, \quad 3)\ \bar{x}(t), u(t) \in Q \cdot \tag{3.17}$$

The proof is identical [2] №7, but in difference from [2] the theorem 3.5 contain the integrability condition.

If $\bar{x}(t), \bar{u}(t) \notin Q$, than $\bar{J}$ is the lower estimation the function (3.12).

If exist the functions *ψ, λ* and at least one pair $\bar{x}(t), \bar{u}(t)$ satisfying (3.17), then any other pair satisfying (3.17) is minimum of the function (3.12) and any allowable minimum the function (3,12) is satisfying the points 1, 2 (3.17) (consicvently remark 3 §1). The set contains the same or better solution, then $\bar{x}(t), \bar{u}(t)$ is

$$N = \{t,x,u : B(t,x,u) + f_0(t,x,u) \geq \bar{B} + \bar{f}_0\} \quad on \quad P^* \times U \text{,}$$

Assume, functions $f_j^i(t,x,u), \varphi^\gamma(t,x,u)$ are continuous and differentiable. Let us take ψ^j in form $\psi^j = p_{ij}(t)x_i$. Let us denote:

$$H = p_{ij}(t) f_{ij}(t,x,u) - f_0(t,x,u) + \lambda_\gamma \varphi^\gamma(t,x,u) \cdot$$

Then p.1 (3.17) of theorem 3.4 we can rewrite: $H(\bar{u}) = \sup_{u \in U} H$ and nessusary condition of minimum (stationarity condition) following from p.2 (3.17) gives:

$$\frac{\partial B}{\partial x_i} = -\frac{\partial p_{ij}}{\partial t_i} - \frac{\partial H}{\partial x_i} = 0, \quad i = 1,2,\ldots,n. \tag{3.18}$$

§4. Inverse substitution method

A. From previous paragraph we have: if we know the minimum any function on acceptable set, we can get information about solution the Problem 1 and solve one from Problem a, b, c, g the §1.

It is known , that the most direct Problems inf $f_0(x)$ on X^* or

$$\inf \int_{t_1}^{t_2} f_0 dt$$

on *Q* (i.e. finding the minimum of main Problem) are difficult or do not have the satisfaction solution. However, if purpose function is not in advance definished, the solution for this non-definished purpse is finding easy. This is not suprising. In mathematics it has long been known that many inverse problems are solved more easily than direct problems. An example, let us consider the problem of finding the roots of an algebraic equation. In the general case for $n > 5$ it is solved with difficulty and her decision (roots) not to be expressed in terms radicals. If the roots are given, then the corresponding algebraic equation may be found easy. On the basis of this idea below it is given method to build function for which an admissible element would be the point of absolute minimum on an admissible set. Since we thus have to solve a problem back to the original problem (not find the minimum given function, but find the function for given the minmum or for fiven field). This method is called **the method of reverse lookup.** The method is presented for two cases: problems of the theory of extrema of functions of a finite numbers of variables (p.B) and optimization problems described by ordinary differential equations (p.C).

B. Let us consider usial Problem of minimum the function of finite variables

$$I = f_0(x), \quad f_i(x) = 0, \quad i = 1,2,...,m < n \,. \qquad (4.1)$$

Let us convert this Problem. Select *m* componets *x* and name them main (base). Suppose for definiteness that this is the first components *m* of the vector *x.* The rest of components $n - m = r$ denote u_j (j = 1,2,...,r).

Rganthe Problem (4.1) we can re-write

$$I = f_0(x,u), \quad f_i(x,u) = 0, \quad i = 1,2,...,m < n \,. \qquad (4.2)$$

where x – m - dimentional vector, $x \in X$, u – r - dimentional vector, $u \in U$.

Let us take more simple purpos function $J_1(x,u)$ and find it's the absolute minimum on $X \times U$. This solution may be used for building of sets *M, N, P*:

$$M = \{x,u : J_1 - f_0 \geq \bar{J}_1 - \bar{f}_0\}, \qquad (4.3)$$
$$N = \{x,u : J_1 + f_0 \leq \bar{J}_1 + \bar{f}_0\}, \qquad (4.4)$$
$$P = \{x,u : J_1 - f_0 \leq \bar{J}_1 - \bar{f}_0\}. \qquad (4.5)$$

Desandvantage this method is next: the some of these sets can di not have the admissible elements (i.e. *x, u* satisfacting f_i = 0).

Assume, the limitations $f_i(x,u) = 0$ in (4.2) may be solved abou *x*:

$$x_i = x_i(u), \quad i = i = 1,2,...,m \qquad (4.6)$$

and $x \in X$ for any $\vee u \in U$.

Assume we take simple function $J_1(x,u)$. Substitute in it's the (4.6) and find $\inf_U J_1(x(u),u), \quad \bar{u}$, and (4.6) $\bar{x}$. This solution is analog (4.3)-(4.5). One may be used for finding sets *M, N, P*. The intersection of these sets with admissible set is not empty. You can take $J_1(x,y,u)$, than $\bar{u} = \bar{u}(y)$. You can use the dependance of *M, N, P* from *y* for changing the "size" of these sets. It is clear assesment

$$\Delta = \inf_{y} \sup_{u} [J_1(x(u),u) - I(x(u)),u] \cdot$$

C. In point 2 §3 we considered the optimization Problem described by conventional differencial equations

$$I = \int_{t_1}^{t_2} f_0(t,x,u)dt, \quad \dot{x} = f_i(t,x,u), \quad i = 1,2,...,n, \quad u \in U. \tag{4.7}$$

We was shown: if we take some function *ψ(t,x)* and find minimum of $\inf_{x,u} B$ in (t1, t2) and $\inf_{x_1,x_2} A$, we get the minimum of Problem 1 or the its lower estimation.

Statement of the Problem. Let us to state the Problem 1 the other way: the find the function which matches the function *ψ(t.x)* and minimum of this function of the admissible set.

Note. Let us note: the offered statement very different from the back problem of variation calculation. In variation calculation, the back problem states next: we have a curve. Find the function, which gives the minimum in this curve. In common case this problem is more difficult then a direct problem.

In our case the minimum curve not given. We find it by given function *ψ(t,x).*

Theorem 4.1. The minimum function corresponding function *ψ(t,x)* is

$$J_1 = \int_{t_1}^{t_2} B_1(t,x)dt = \int_{t_1}^{t_2} - \inf_{u \in U} [-\psi_{x_i} f_i(t,x,u) - \psi_t]dt \quad . \tag{4.8}$$

And correcponding to it the minimum curve is given by equations

$$\dot{x}_i = f_i[t,x,\bar{u}(t.x,\psi_{x_i}\psi_t)], \quad i = 1,2,...,n, \tag{4.9}$$

where $\bar{u} = \bar{u}(t,x,\psi_{x_i},\psi_t)$ we find from (4.8).

Proof. Write the expression *B* (see (3.11)) for problem (4.7) and check up condition (3.8) of theorem 3.1:

$$B_2(t) \equiv \inf_{x,u} \inf [B_2(t,x) - \psi_{x_i} f_i(t,x,u) - \psi_t] \cdot \tag{4.10}$$

Obviosly, the (4.10) identically equals zero for *ψ* = *ψ(t,x*) from (4.8) and $\bar{x}, \bar{u}$ satisfacting (4.7). If we take as *x(t₂)* the value *x(t)*, received from (4.9) for t_2, then the point 2 (3.8) disappear and all condition (3.8) of theorem is executed. Theorem is prooved.

Consequence. If $B_1 = f_0(t,x)$, then x(t) getting from (4.10) give the set of the minimal curves for boundary condition $\psi_2 = \psi$. In particulary, if the end of curve *x(t)* from (4.9) match with given boundary conditions, that this curve is minimum curve of Problem 1.

Note. Boundary conditions in the left end can always be performed. For it we must start the intgration from the given conditions (4.9). We can perform the boundary condition in the right end the next method. Take in form *ψ(t,x,c)* where *c – n* – dimentional constant. Substitute *ψ(t,x,c)* in (4.9) and select *c* such that to perform the given end condition in the right end.

Getting numerical function may be used for receiving the set *N, P* of Theorem 3.3 :

$$N-\{t,x:\ f_0+B_1\le \bar f_0+\overline{B}_1\},\quad P=\{t,x:B_1-f_0\le \overline{B}_1-\bar f_0\},$$

$$\text{where}\quad f_0=f_0[t,x,\bar u(t,x,\psi_x,\psi_t)],\quad \psi(t,x)\quad \text{is}\quad \text{given.}$$

If we find

$$\bar J=\psi_2-\psi_1+\int_{t_1}^{t_2}\inf_x(f_0-B_1)dt\ ,$$

We get also the lower estimation.

Memo, the assignment *ψ(t,x)* gives us not single nometical function and its point of minimum. One gives a set of minimums satisfaction the boundary conditions $\psi_2-\psi_1=c$.

Note: We can take *ψ(t,x,y)*. Then $B_1(t,x,y)$. If we can select such $\bar y(t)$ that $B_1(t,x,\bar y)=f_0(t,x)$ and boundary conditions is perfomed, then $\bar u(t,x,\bar y)$ is the optimal synthesis of Problem 1.

D. We also show: how you can find the numerical function for given the syntes of control *u = u(t,x).*

Equate the given *u = u(t,x)* to the control findedfrom (4.8). We get the equation in particular derivities

$$u(t,x)=\bar u(t,x,\psi_{x_i},\psi_t)\,. \tag{4.11}$$

Substitute its solution *ψ(t,x)* and given u(t,x) in (4.8), we find the numerical corresponding function. If $B_1=f_0(t,x)$ that is synthesis the Problem 1 for the bounded condition $\psi_2=\psi$.

Possible the other method . We take *u = u(t,x,y).* Substitute its in (4.8). Then $B_1=B_1(t,x,c,y)$. We can try using *y* to reach the identify $f_0\equiv B_1$ and using *c* to minimize the nymerical function *I*.

Example 4.1. Let us consider the task of design the regulator

$$I=\int_{t_1}^{t_2}b_{ij}x_ix_jdt, \tag{4.12}$$

$$\dot x_i=a_{ij}x_j+u,\quad 0\le t\le\infty, \tag{4.13}$$

$$x_i(0)=x_{i,o},\quad x_i(\infty)=0, \tag{4.14}$$

where $f_0=b_{ij}x_ix_j$ is the positive definite form.

Take $u=c_ix_i$, where c_i are constants. Let us to search *ψ* as the quadratic form $\psi=A_{ij}x_ix_j$ with unknown coefficients. Equate $f_0\equiv\dot\psi$:

$$b_{ij}x_ix_j=A_{ij}x_i(a_{ij}x_j+c_jx_j).$$

Let us equate coefficient in same x_i, x_j in left and right of this equation. We get the set *n*(*n*+1)/2 the linear inhomogenius equations having the same number of unknown A_{ij}. If the determinant of this system $\Delta \neq 0$, we find A_{ij}. We substitute $f_0 \equiv \dot{\psi}$ in (4.12), integrate and find $I = \psi(\infty, c) - \psi(0, c)$ or using (4.14) $I = -\psi(x_{io}, c)$. When we find minimum of this expression for *c*,we get the optimal syntes.If $-\psi(x, \bar{c})$ is the positive definite form then this function is the Lyapunov function (because $-\dot{\psi} \geq 0$ and the regulator is assimptotic stable.

§5. Method of combining extrema in problems of constrained minimum.

We will show in this paragraph that method combining extrema, considered in §2 the Chapter 1, it is apply in tasks of theory the functions of a finite number of variables (point A) and tasks described the conventional difference equations.

A) Let us again consider the Problem of the theory the functions of a finite number of variables

$$I = f_0(x), \quad f_i(x) = 0, \quad i = 1,2,\ldots, m. \tag{5.1}$$

Write the numerical function

$$J(x,c) = f_0(x) + \beta(x,c) + \alpha_1(x), \tag{5.2}$$

Here $\alpha_1(x)$ is α – function, *c* is *n* – dimentional constant.

From condition

$$\inf_{x \in X^*} J(x,c), \tag{5.3}$$

we find $\varphi_1(x^{(1)}, c) = 0.$

From condition

$$\Phi(x,c) = \sup_{x \in X^*} [\beta(x,c) + \alpha_2(x)], \tag{5.4}$$

we find $\varphi_2(x^{(2)}, c) = 0.$ Solve equations φ_1, φ_2 together with (5.1) (cjmbining equations):

$$\varphi_1(x^{(1)}, c) = 0, \quad \varphi_2(x^{(2)}, c) = 0, \quad x^{(1)} = x^{(2)}, \tag{5.5}$$

we receive the absolite minimum the Problem 1. The eddittive $\beta(x,c)$ selectes so that tasks (5.3), (5.4) are solved easier.

For example, $\alpha_1 = \lambda_i f_i, \quad \alpha_2 = \nu_i f_i.$ Functions $f_i(x)$, i = 0,1, …,n are continious and difference , the functions $J(x,c)$, $\Phi(x,c)$ have single minimum and maximum for any *c*. That we have system (3*n* + 2*m*) equations with same numbers of unknown magnitudes α(1), α(2), c, λ, ν.

Example is not include.

B) Let us to consider the task, described the conventional different equations:

$$I=\int_{t_1}^{t_2} f_0(t,x,u)dt, \quad \dot{x}_i = f_i(t,x,u), \quad i=1,2,...,n, \quad u\in U, \quad x(t_1)=x_1, \quad x(t_2)=x_2 \text{ ,} \tag{5.9}$$

Take ψ in form $\psi^{(1)} = p_i^{(1)}(t)\,\alpha_i^{(1)}$ and create the function

$$B_1 = f_0 + \beta(t,x^{(1)},u^{(1)},z) - p_i^{(1)} f_i^{(1)} - \dot{p}_i^{(1)} x_i^{(1)} = -H^{(1)} - \dot{p}_i^{(1)} x_i^{(1)} \cdot$$

Here *z(t)* is *r* – dimentinal function. One can have the limited gaps the first type.

From $\inf\limits_{x,u} B_1$ and (5.9) we find

$$\dot{p}^{(1)} = -H_x^{(1)}, \quad \bar{u}^{(1)} = \bar{u}^{(1)}(t,x^{(1)},p^{(1)},z), \quad \dot{x}^{(1)} = f(t,x^{(1)},u^{(1)}). \tag{5.10}$$

Take $\psi^{(2)} = p_i^{(2)} x_i^{(2)}$ and create the function

$$B_2 = \beta(t,x^{(2)},u^{(2)},z) - p_i^{(2)} f_i^{(2)} - \dot{p}_i^{(2)} x_i^{(2)} = -H^{(2)} - \dot{p}_i^{(2)} x_i^{(2)} \cdot$$

From $\inf\limits_{x,u} B_2$ and (5.9) we find

$$\dot{p}^{(2)} = -H_x^{(2)}, \quad \bar{u}^{(2)} = \bar{u}^{(2)}(t,x^{(2)},p^{(2)},z), \quad \dot{x}^{(2)} = f(t,x^{(2)},u^{(2)}). \tag{5.11}$$

Using the combining equation: $x^{(1)} = x^{(2)}, \quad u^{(1)} = u^{(2)}$ we get final:

$$\dot{x} = f(t,x,u^{(1)}), \quad \dot{p}^{(1)} = -H_x^{(1)}, \quad \dot{p}^{(2)} = -H_x^{(2)}, \quad \bar{u}^{(1)}(t,x,p^{(1)},z) = \bar{u}^{(2)}(t,x,p^{(2)},z) \text{ ,} \tag{5.12}$$

That is system 3*n + r* equations with 3*n + r* unknown *x*, $p^{(1)}$, $p^{(2)}$, *z* . Last equation in (5.12) is the combining equation. The additive function *β* selecting so that the solution task of finding *inf* and *sup* were simpler.

§6. Generalizing the Theorem 3.1 in case the brocken *ψ(t,x)*.

Theorem 6.1. *Assume there is numerical function ψ(t,x) defined on set* $T\times G$ *, bounded below, piecewise differentiable and piecewise continuous. The function ψ(t,x) and its derivatives can have the breaks the first types on the limited set* $\Phi_s(t_s,x),\ s=1,2,...,k-1$ *zero measure. This function is such that there is:*

$$1)\ \inf_R\,(F+\psi_k-\psi_o), \quad 2)\ \inf_{t_s,x\in\Phi_s}(\psi_s^- - \psi_s^+), \quad \bar{t}_s \succ \bar{t}_{s-1}, \quad t_k^{'} \succ \bar{t}_{k-1}^{'}, \quad s=1,2,...,k-1,$$
$$3)\ \inf_{G\times T} B = 0, \quad 4)\ \bar{x}(t), \bar{u}(t) \in Q.$$

Then $\bar{x},\bar{u}$ *(are got from points 1 -3) is the absolute minimum the Problem 1.*

Here ψ_s^-, ψ_s^+ are value *ψ* in left and right side (along $\bar{x}(t)$) of the breaks the function *ψ* and its derivatives.

Proof: From poins 1 – 3 we have

$$\bar{J} = \inf_{R}(F + \psi_k - \psi_0) + \sum_{s=1}^{k-1} \inf_{t_s, x}(\psi_s^- - \psi_s^+) + \sum_{s=0}^{k-1} \int_{t_s}^{t_{s-1}} \inf_{x,u} B dt \,.$$

On feasible cirves (from *Q*) the $\bar{J}$ convert in function $I = F + \int_{t_1}^{t_2} f_0 dt$. In this case if we apply the consequence 4 , §1, point 4 of the theorem statement is obviously. Theorem is prooved.

Note. The conditions 3 of Therem 6.1 is sometimes difficult to check up. In this case the requirements 2 - 3 of theorem 6.1 we can change the damage

$$\inf_{t_s}[\inf_{x}(\psi_s^- - \psi_s^+) + \int_{s-1}^{s} \inf_{G\times U} B dt + \int_{s}^{s+1} \inf_{G\times U} B dt] \,.$$

One must be checked up in every point t_s, s = 1,2,…,k-1.

§7. Optimization the problems described the conventional differential equations having the limitations.

We find minimum *A, B* in Theorem 3.1, chapter II on the corresponding sets *R* and $U \times G$. The most widely method of separating the feasible sets is the separation of them from more widely set by equilities and inequilities. In this case, we can solve our problem by the methods the *α*- and *β*-functions.

Let us shortly consider the most common cases.

1. Limitations are the equilities

a) Assume the admissible set *R* is separated by equilities:

$$g_i(x_1, x_2) = 0, \quad i = 1,2,..., l < 2n \,. \qquad (7.1)$$

Then the task *inf A* we can change the task

$$\inf_{x_1, x_2}[A + \mu_i(x_1, x_2, z_i) g_i(x_1, x_2)] \,. \qquad (7.2)$$

Here μ_i is known functins, *z* is *l*-dimentional unknown vector. In particulary, we can take $\mu_i = z_i$.

b) Assime the admissible set $U \times G$ is separated by equilities

$$\varphi_i(t, x, u) = 0, \quad i = 1,2,..., l < r \,. \qquad (7.3)$$

Assume, we can find from (7.3) the *l* component the vector *u*. Than the problem $\inf\limits_{G\times U} B$ we can change the problem

$$\inf_{x,u}\,[B+\lambda_i(t,x,w)\varphi_i(t,x,u)]\,, \tag{7.4}$$

Where λ_i are known function, w_i is *l* - dimantional unknown vector function. In particular, we can take $\lambda_i = w_i$.

c) Assume the admisseble set G is separated by the equilities

$$\varphi_i(t,x)=0,\quad i=1,2,...,l<r\,. \tag{7.5}$$

Differenciate (7.5) full case for *t* and find

$$\varphi_i^{(1)}(t,x,u)\equiv\frac{\partial\varphi_i}{\partial x_j}f_j(t,x,u)+\frac{\partial\varphi_i}{\partial t}=0,\quad i=1,2,...,l<n\,. \tag{7.6}$$

If in system (7.6) there is equations do not contain *u*, we differenciate them next time and so on whole we get the the system where all *l* eqution contain u. Assume we can find all *l* components from this system ($l < r$).

Than the problem (7.5) is redused to the tasks the point *a, b* in which (7.6) is (7.3), but (7.5) and all equtions (7.6) not contain *u*, are (7.1).

2. Limitations are inequalities. (excerpt)

a) Feasible set *R* is allocated by inequalities:

$$g_i(x_1,x_2)\le 0,\quad i=1,2,...,l.$$

Then acording the Teorem 1.4 Chapter 1 we change the problem $\inf\limits_{R} A$ by problem (7.2) with the additional conditions:

$$\bar\lambda_i\bar g_i=0,\quad \bar\lambda_i\ge 0\quad (\text{here}\quad i\quad \text{is}\quad \text{not}\quad \text{sum}) \tag{7.7}$$

b) Feasible set $U\times G$ is allocated by inequalities:

$$\varphi_i(t,x,u)\le 0,\quad i=1,2,...,l. \tag{7.8}$$

All inequalities contain *u.* Then the task $\inf\limits_{U\times G} B$ we change the task (7.4) wuth conditions

$$\bar\lambda_i\bar\varphi_i=0,\quad \bar\lambda_i\ge 0\quad (\text{here}\quad i\quad \text{is}\quad \text{not}\quad \text{sum}) \tag{7.9}$$

Example 7.1. Assume in task

$$I = \int_{t_1}^{t_2} f_0(t,x,u)dt, \quad \dot{x}_i = f_i(t,x,u), \quad i = 1,2,...,n,$$

Control *u* is scalar, the feasible set *U* limited inaquility $a \le u \le b, \ (a < b)$. Compose (7.4):

$$\inf_U [B + \lambda_1(u-b) + \lambda_2(-u+a)].$$

According (7.9) on feasible u: $\bar{\lambda}_1(\bar{u}-b) = 0, \quad \bar{\lambda}_2(-\bar{u}+a) = 0.$ That way we have

$$\inf_u [B + \lambda_1(u-b) + \lambda_2(-u+a)] = \inf_{u_1,u_2 \in U} B.$$

In right side we have one condition the Pontryagin method.

(Part of the text are missing)

==

…….

==

§10. Note on the equivalence of different forms of variational problems

A) In §3 the next problem of minimization was considered

$$I = F(x_1,x_2) + \int_{t_1}^{t_2} f_0(t,x,u)dt, \tag{10.1}$$

on solution of equations

$$\dot{x}_i = f_i(t,x,u), \quad i = 1,2,...,n. \tag{10.2}$$

In the theoretical analysis for the sake of simplicity, we often assume that in (3.1)
$F \equiv 0 \quad or \quad f_0 \equiv 0.$

We show that it does not restrict the generality of our reasoning.

Take

$$I = \int_{t_1}^{t_2} f_0(t,x,u)dt$$

And differentiate it for the variable upper limite *t* and designate $\dot{x}_{n+1} = f_0$. We get the task

$$I = x_{n+1}(t_1), \quad \dot{x}_i = f_i, \quad \dot{x}_{n+1} = f_0. \tag{10.3}$$

B) Assume $I = F(x_1,x_2).$ Differenciate it by *t* and integrate, we get numerial function

$$I = \int_{t_1}^{t_2} (F_{x_i} f_i)dt \tag{10.4}$$

We can same way to convert (10.1) in (10.4) and in (10.3).

C**)** Let us to assume the (10.1) and (10.2) depend from constants c_k which must be optimal. Designate ck = xi+k and add to (3.3) equation $\dot{x}_{n+k} = 0$. We reduced the task having the optimising constants to conventional task.

In practice it is camfortable to solve the problema (10.1), (10.2) with constant parameters. Than to change them (for example the gradient method) so, the function (3.1) decreases.

D) The problem with $f_i(t,x,u)$ which obviously depend from *t* , we can reduse to problem $f_i(x,u)$ do not depend obviously from *t* , if to designate $t = x_{n+1}$ and add to (10.1) the equation $\dot{x}_{n+1} = 1$.

C) Let us to show how the task with the mooving ends t_1 and t_2 we can reduse the task with fix interval of integrate. Take the new variable $t = c\tau$. Than task (10.1),(10.2) having variables t1 or t2 was redused in task with fix interval (τ_1,τ_2):

$$I = F + \int_{t_1}^{t_2} cf_0(\tau, x, u)d\tau, \quad x' = cf_i(c\tau, x, u),$$

where the touch means the derivative for τ. The constant $c > 0$ is selected from minimum I.

Application to Chapter II.

1. Theorem 3.1 and known methods of solution the problem described the ordinary differential equations.

From Theorem 3.1 we can to get the conditions which are same with known algorithms of optimal control, for example: Pontriagin principle [1], Bellman equation [6], classicад calculus of vatiation [7],

Let us to request additional that function *f, ψ* have the need continious derivatives.

a) Pontriagin principle. According [2] take $\psi(t,x)$ in form $\psi = p_i(t)\Delta x_i$, where $p_i(t)$ are some differenciable functions *t*, $\Delta x_i = x_i - \bar{x}_i$.Create the Hamiltonian

$$H = p_i f_i(t,x,u) - f_0(t,x,u). \tag{1}$$

Then $B = - H - p_i x_i$. Necessary condition of the minimum *B* for *x*, which follows from p.1 (3.8) of Theorem 3.1 (stationarity condition) is

$$B_{x_i} \equiv -p_i - H_{x_i} = 0, i = 1,2,...,n. \tag{2}$$

Moreover of claim 1 (3.8) we have

$$B(t,x,\bar{u}) = \inf_{u \in U} B(t,x,u) \quad \text{or} \quad \inf_{u \in U}(-H) = -\sup_{u \in U} H \qquad (3)$$

Terms and conditions (2), (3) together with (3.3) coincide with the corresponding terms and conditions of the Maximum principle* [1].

b) Belman equation. Assume $x_n \neq 0$. Take all $\lambda_i = 0$ $i = 1,2,\ldots,n-1$ with exeption $\lambda_n = \psi(t,x)/x_n$. Sabstitute them in (3.9) §3, we get the known Belman equation [6]

$$\inf_{u \in U}(f_0 - \psi_{x_i} f_i - \psi_t) = 0 \qquad (4)$$

Boundary condition for them is *A = const*. Solution of this equation is the field of all optimal trajectories.

c) Classical calculus of variation. From claims 1, 2 Theorem 3.1 easy to get the conditions of a relative minimum coinciding with the relevant terms of the calculus of variations [7].

Let us assune *U* is the open area, $\dot{x}(t), u(t)$ are continiosly, $f_i(t,xu)$ have continious partial deriveties up the third order. Take $\psi = p_i(t)\Delta x_i$. From (3) that at minimum

$$B_{u_i}(t,x,u) = -H_{u_i}(t,x,u) = 0, \quad i = 1,2,..,r\,, \qquad (5)$$

Equtions (2),(4) equal the conventional Eiler-Lagrange equations [7] §2 p.1. From [3] also follow

$$-H_{u_i u_j}\delta u_i \delta u_j \geq 0, \quad i,j = 1,2,\ldots,r\,. \qquad (6)$$

That matches with Klebs condition.

(Itanslation of the Chapter 2 is not finished)

===

………

==

Attachment 1

Chapter 4A from book "Non-Rocket Space Launch and Flight", by Alexander Bolonkin, Elsevier, 2007. 488 pgs. ISBN-13: 978-0-08044-731-5, ISBN-10: 0-080-44731-7 . pp.383-422

Optimal Trajectories of Air and Space Vehicles

Summary

The author has developed a theory on optimal trajectories for air vehicles with variable wing areas and with conventional wings. He applied a new theory of singular optimal solutions and obtained in many cases the optimal flight. The wing drag of a variable area wing does not depend on air speed and air density. At first glance the results may seem strange, however, this is the case and this chapter will show how the new theory may be used. The equations that follow enable computations of the optimal control and optimal trajectories of subsonic aircraft with pistons, jets, and rocket engines, supersonic aircraft, winged bombs with and without engines, hypersonic warheads, and missiles with wings.

The main idea of the research is to use the vehicle's kinetic energy to increase the range of missiles and projectiles.

The author shows that the range of a ballistic warhead can be increased 3–4 times if an optimal wing is added to it, especially a wing with variable area. If we do not need increased range, the head mass of rockets can be increased. The range of large gun shells can also be increased 3–9 times. The range of an aircraft may be improved by 3–15% or more.

The results can be used for the design of aircraft, space ship, head of rockets, missiles, flying apparatus and shells for large guns.

Key words: Methods of optimizatiom, optimization, optimal control, aviation, space ships.

Nomenclature (in metric system)

a – the speed of sound, m/s,

a_1, b_1, a_2, b_2 – coefficients of exponential atmosphere,

C_L – lift coefficient,

C_D – drag coefficient,

C_{Do} – drag coefficient for $C_L = 0$,

C_{DW} – wave wing drag coefficient when $\alpha = 0$,

C_{Db} – body drag coefficient,

c – relative thickness of a wing,

c_b – relative thickness of a body,

c_1 – relative thickness of a vehicle body,

c_s – fuel consumption, kg/s/ kg thrust,

$\overline{D}$ – drag of vehicle, N,

D – drag of vehicle without α, N,

D_{0W} – wave wing drag when $\alpha = 0$, N,

D_{0b} – drag of a vehicle body, N,

H – Hamiltonian,

h – altitude, m,

$K = C_L/C_D$ – the wing efficiency coefficient,

k_1, k_2, k_3 – vehicle average aerodynamic efficiencies for sub-distances 1, 2, 3 respectively,

L – range,

$M = V/a$ – Mach number,

m – mass of vehicle, kg,

$p = m/S$ – load on a square meter of wing,

$q=\rho V^2/2$ – a dynamic air pressure,

R – aircraft range or R = distance from flight vehicle to Earth center;
$R = R_o + h$, where R_o = 6378 km is Earth radius,

t– time,

$T = V_e\beta$ – thrust, N,

V – vehicle speed, m/s,

V_e – speed of throw back mass (air for propeller engine, jet for jet and rocket engine), m/s,

S – wing area, m^2,

s – length of trajectory,

T – engine thrust, N,

Y – lift force, N,

α – wing attack angle,

β – fuel consumption,

θ – angle between the vehicle velocity and the horizon,

ω – thrust angle between thrust and velocity,

ω_E – Earth angle speed,

φ_E – lesser angle between the Earth's Polar axis and a perpendicular to a flight plate,

ρ – air density. kg/m^3.

Introduction

The topic of the optimal flight of air vehicles is very important. There are numerous articles and books about the optimal trajectories of rockets, missiles, and aircraft. The classical research of this topic is by Miele[1]. Unfortunately, the optimal theory of this problem is very complex. In most cases, the researchers obtained complex equations, that allow one to compute a single optimal trajectory for a given aircraft and for given conditions, but the structure of optimal flight is not clear and simple formulas of optimal control (which depend only on flight conditions) are absent.

The author's new theory of singular optimal solutions, developed earlier[2–14], does not contain unknown coefficients or variables as previous theories have. He found that the optimal flight path depends only on the flight conditions and the addition of certain variable wing structures.

In conclusion, the author applies his solution to ballistic missiles, warheads, flying bombs, large gun shells, and subsonic, supersonic, and hypersonic aircraft with rocket, turbo-jet, and propeller engines. He shows that the range of these air vehicles can be increased 3–9 times.

1. General equations

Let us consider the movement of an air vehicle given the following conditions: (1) The vehicle moves in a plane containing the Earth's center. (2) The vehicle design allows the wing area to be changed (this will prove important in the remainder of this chapter). (3) We ignore the centrifugal force from the Earth's rotation (it is less then 1%). (4) Earth has a curvature.

Then the equations for flying vehicle (in a system of coordinates where the center of the system is located at the center of gravity of the flying vehicle, the *x*-axis is in the direction of flight, the *y*-axis is perpendicular to the *x*- axis, Fig. A4.1) are

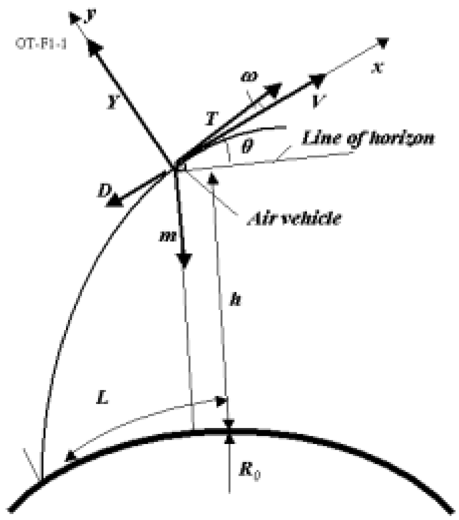


Fig. A4.1 Vehicle forces and coordinate system.

$$\frac{dL}{dt} = V\cos\theta \quad ,$$
$$\frac{dh}{dt} = V\sin\theta \quad , \qquad \text{(A4.1)–(A4.2)}$$

$$\frac{dV}{dt} = \frac{T(h,V,\beta)\cos\omega - \overline{D}(\alpha,V,h)}{m} - g\sin\theta\,,$$
$$\frac{d\theta}{dt} = \frac{T(h,V,\beta)\sin\omega + Y(\alpha,V,h)}{mV} - \frac{g}{V}\cos\theta + \frac{V\cos\theta}{R} + 2\omega_E\cos\varphi_E\,, \qquad \text{(A4.3) – (A4.5)}$$
$$\frac{dm}{dt} = -\beta \quad .$$

All values are in the metric system and all angles are taken to be in radians.

Flight with a small change of vehicle mass and flight path angle

Most air vehicles fly at an angle θ in the range $\pm 15^{\circ}$ ($\theta = \pm 0.2618$ rad), with the engine located along the velocity vector. This means

$$\sin\theta = \theta, \qquad \cos\theta = 1\,, \qquad \omega = 0\,, \qquad \text{(A4.6) – (A4.8)}$$

because $\sin 15^{\circ} = 0.25882$, $\cos 15^{\circ} = 0.9659$.

Let as substitute (A4.6) – (A4.8) into (A4.1) – (A4.5)

$$\frac{dL}{dt} = V \quad , \qquad \frac{dh}{dt} = V\theta \quad , \tag{A4.9) – (A4.10}$$

$$\frac{dV}{dt} = \frac{T(h,V) - \overline{D}(\alpha,V,h)}{m} - g\theta \, , \qquad \frac{d\theta}{dt} = \frac{Y(\alpha,V,h)}{mV} - \frac{g}{V} + \frac{V}{R} + 2\omega_E \cos\varphi_E \, , \tag{A4.11) – (A4.12}$$

$$\frac{dm}{dt} = -\beta \quad , \tag{A4.13}$$

where

$$|\theta| \le \theta_{\max} \, . \tag{A4.14}$$

Many air vehicles fly with a low angular speed of $d\theta/dt$. The change of mass is also low in flight. This means m = const, $dm/dt \cong 0$.

$$d\theta/dt \approx 0, \qquad dm/dt = 0 \, . \tag{A4.15) – (A4.16}$$

Let us take a new independent variable s = length of trajectory

$$dt = ds/V, \tag{A4.17}$$

and substitute (A4.14)-(A4.17) in (A4.9)-(A4.13). Then system (A4.9)-(A4.13) takes the form

$$\frac{dL}{ds} = 1 \, , \qquad \frac{dh}{ds} = \theta \, , \qquad \frac{dV}{ds} = \frac{T(h,V) - \overline{D}(\alpha,V,h)}{mV} - \frac{g}{V}\theta \, , \qquad 0 = \frac{Y(\alpha,V,h)}{mV} - \frac{g}{V} + \frac{V}{R} + 2\omega_E \cos\varphi_E \, . \tag{A4.18) – (A4.21}$$

Let us re-write equation (A4.21) in the form

$$Y(\alpha,V,h) - mg + \frac{mV^2}{R} + 2mV\omega_E \cos\varphi_E = 0 \, . \tag{A4.22}$$

If we ignore the last element, equation (A4.22) takes the form

$$Y(\alpha, V, h) - mg + \frac{mV^2}{R} = 0 \ . \qquad \text{(A4.22)'}$$

If *V* is not very large (*V* < 3 km/s), the two last elements in equation (A4.21) are small and they may be ignored. Equations (A4.22) and (A4.22)' can be used for deleting α from $\overline{D}$.

Note the new drag without α is

$$D = D(h, V). \qquad \text{(A4.23)}$$

If we substitute α from (A4.22) into equation (A4.20) the equation system take the form

$$\frac{dL}{ds} = 1,$$
$$\frac{dh}{ds} = \theta, \qquad \text{(A4.24) – (A4.26)}$$
$$\frac{dV}{ds} = \frac{T(h, V) - D(V, h)}{mV} - \frac{g}{V}\theta,$$

Here the variable θ is new control limited by

$$|\theta| \le \theta_{\max}. \qquad \text{(A4.27)}$$

Statement of the problem

Consider the problem: finding the maximum range of an air vehicle described by equations (A4.24) – (A4.26) for the limitation (A4.27). This problem may be solved using conventional methods. However, it is a non-linear problem but contains the linear control, which means the problem has a singular solution. To find this singular solution, we will use methods developed previously[2, 4].

Write the Hamiltonian (for purpose – minimum of time):

$$H = 1 + \lambda_1\theta + \lambda_2 \frac{1}{V}\left(\frac{T - D}{m} - g\theta\right), \qquad \text{(A4.28)}$$

where $\lambda_1(s), \lambda_2(s)$ are unknown multipliers. Application of the conventional method gives

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial h} = -\lambda_2 \frac{1}{V}\left(\frac{T_h' - D_h'}{m}\right) \quad ,$$
$$\dot{\lambda}_2 = -\frac{\partial H}{\partial V} = -\lambda_2\left[-\frac{1}{V^2}\left(\frac{T-D}{m} - g\theta\right) + \frac{1}{V}\left(\frac{T_V' - D_V'}{m}\right)\right] \quad , \qquad \text{(A24.29) – (A4.31)}$$
$$\theta = \max_{\theta} H = \theta_{\max} sign\left[\lambda_1 - \lambda_2 \frac{g}{V}\right] \quad .$$

Where $D_h^{'}$, $D_v^{'}$, $T_h^{'}$, $T_v^{'}$ denote the first partial derivatives of *D, T* by *h*, *V* respectively.

The last equation shows that the control θ can have only two values $\pm\theta_{\max}$. We consider the singular case when

$$A = \lambda_1 - \lambda_2 \frac{g}{V} \equiv 0 \quad . \qquad \text{(A4.32)}$$

This equation has two unknown variables λ_1 and λ_2 and does not contain information about the control θ. Let us to differentiate equation (A4.32) for the independent variable *s*. After substitution the equations (A4.26), (A4.29), (A4.30), and (A4.32) into the result of differentiation , we obtain the relation for $\lambda_1 \neq 0$, $\lambda_2 \neq 0$

$$V\left(T_h' - D_h'\right) = g\left(T_V' - D_V'\right) \qquad \text{(A4.33)}$$

This equation does not contain θ either, but it contains the important relation between the variables *V* and *h* on the optimal trajectory.

If we have the formulas (or graphs)

$$D = D(h,V), \qquad \text{(A5.34)}$$

$$T = T(h,V), \qquad \text{(A4.35)}$$

we could find the relation

$$h = h(V) \qquad \text{(A4.36)}$$

and the optimal trajectory for a given air vehicle.

This also gives important information about the structure of the optimal solution. Investigation of equation (A4.33) shows that the equation has one solution in each of the subsonic, supersonic, and hypersonic fields. The equation can have two solutions for a transonic field.

This means the optimal trajectory in most cases has three parts (see Fig. A4.2):

a) When climbing and in flight a vehicle moves from the initial point *A* with the angle $\pm\theta_{\max}$ up to the optimal curve (A4.36), then continues along the optimal curve (A4.36) and moves with at an angle $\pm\theta_{\max}$ to point *B*.

b) When descending and in flight (Fig. A4.3) a vehicle moves from the initial point *A* with the angle $\pm\theta_{\max}$ (up or down) to the optimal curve (A4.36), then continues down the optimal curve (A4.36), and moves at an angle $\pm\theta_{\max}$ (up or down) to the point *B*.

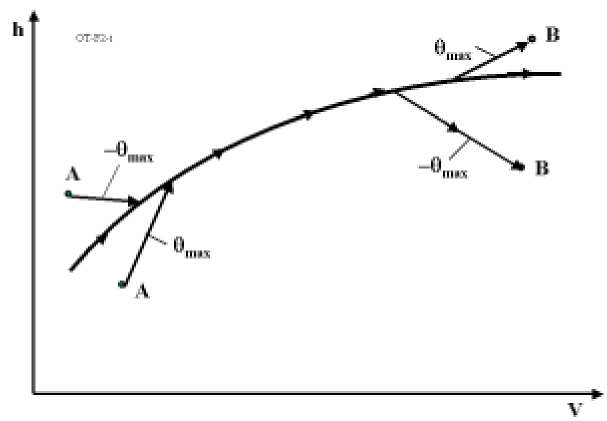


Fig. A4.2.. Optimal trajectory for air vehicle climb and flight.

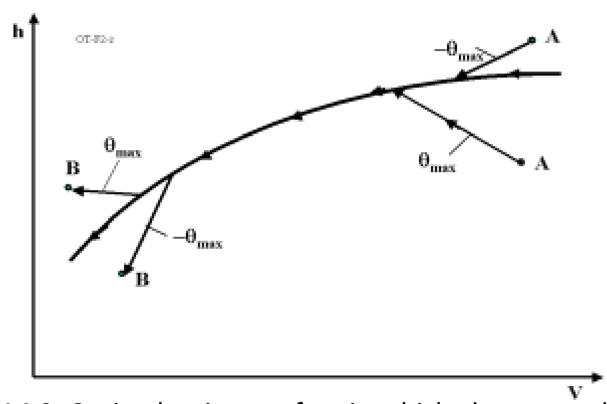


Fig. A4.3. Optimal trajectory for air vehicle descent and flight.

The selection of direction (up or down, with $\theta_{\max}$ or $-\theta_{\max}$ respectively) depends only on the position of the initial and end points *A* and *B*.

For air vehicles with rocket engines *T* = const, equation (A4.33) has a very simple form

$$VD'_h = gD'_V \ . \qquad \text{(A4.37)}$$

The same form (same curve) also applies for a ballistic warhead, which does not have engine thrust (after its short initial burn) (*T* = 0).

If we want to find an equation for the control θ, we continue to differentiate equation (A4.33) with the independent variable *s*, and substitute into the equations (A4.25), (A4.26), (A4.29), (A4.30), (A4.32), and (A4.33). We obtain the relation for θ if $\lambda_1 \neq 0$, $\lambda_2 \neq 0$

$$\theta = \frac{B_1(T-D)}{mV\left(B_1 \frac{g}{V} - B_2\right)} \quad , \qquad \text{(A4.38)}$$

where

$$B_1 = (T'_h - D'_h) + V(T''_{hV} - D''_{hV}) - g(T''_{VV} - D''_{VV}) \quad ,$$
$$B_2 = V(T''_{hh} - D''_{hh}) - g(T''_{hV} - D''_{hV}) \quad . \qquad \text{(A4.39)–(A4.40)}$$

Here signs in form D''_{hV} are the second partial derivates *D* for *h, V*.

$$D''_{hV} = \frac{\partial^2 D}{\partial h \partial V} \quad . \qquad \text{(A4.41)}$$

If the thrust does not depend on *h, V* (*T* = const) or no engine (*T* = 0), the equation for θ becames simpler

$$\theta = \frac{[(gD''_{VV} - D'_h) - VD''_{hV}](T-D)}{m[g(gD''_{VV} - D'_h) + V^2 D''_{hh}} \, . \qquad \text{(A4.42)}$$

In accordance with other publications[2–8] (e,g, equation (4.2)[4]) the necessary condition for optimal trajectory is

$$-(-1)^k \frac{\partial}{\partial \theta}\left[\frac{d^{2k}}{ds^{2k}}\left(\frac{\partial H}{\partial \theta}\right)\right] \geq 0 \, . \qquad \text{(A4.43)}$$

where *k* = 1.

To obtain results for different forms of the drags and thrusts, we must take formulas (or graphs) for subsonic, transonic, supersonic, or hypersonic speed, and specific formulas for the thrust and substitute them in the equation (A4.33) and (A4.38). Consider two cases: subsonic and hypersonic speeds.

Subsonic speed (*V* < 270 m/s) and different engines.

Lift, drag, and derivative equations for subsonic speed are

$$L = mg = \varsigma\alpha\frac{\rho V^2}{2}S\,,\quad \overline{D} = C_D\frac{\rho V^2}{2}S\,,\quad C_D = C_{Do} + \varepsilon\alpha^2\,,\quad D = \left[C_{Do} + \varepsilon\left(\frac{2mg}{\varsigma\rho V^2 S}\right)^2\right]\frac{\rho V^2}{2}S\,,$$

$$\rho = a_1 e^{-h/b_1}\,,\quad \frac{\partial D}{\partial V} = \left[C_{Do} - \varepsilon\left(\frac{2mg}{\varsigma\rho V^2 S}\right)^2\right]\rho VS\,,\quad \frac{\partial D}{\partial h} = -\frac{1}{b_1}\left[C_{Do} - \varepsilon\left(\frac{2mg}{\varsigma\rho V^2 S}\right)^2\right]\frac{\rho V^2}{2}S\,,$$

(A4.44)

where $\varsigma = \frac{6.24\lambda}{\lambda+2}$, $\varepsilon = \frac{\varsigma^2}{\pi\lambda}$, magnitude $\varepsilon \approx \zeta^2/\pi\lambda$ is an induced drag coefficient, $\lambda = l^2/S$, l is a wing span.

It is known in conventional aerodynamics that the coefficient of flight efficiency k is

$$k = \frac{C_L}{C_D} = \frac{\varsigma\alpha}{C_{Do} + \varepsilon\alpha^2}\,,\quad from\quad \max_{\alpha} k\quad we\ obtain\quad \alpha_{opt} = \sqrt{\frac{C_{Do}}{\varepsilon}}\,,\quad k_{\max} = \frac{\varsigma}{2\sqrt{\varepsilon C_{Do}}}\,. \quad \text{(A4.45)}$$

a) Aircraft with rocket engine. For this aircraft the thrust T is constant or *0*. Equation (A4.33) has form (A4.37). Find the partial derivatives

$$T'_V = 0\,,\quad T'_h = 0\,. \quad \text{(A4.46)}$$

Substituting (A4.44) to (A4.46) in (A4.37) we obtain the relation between air density ρ, altitude h, and aircraft speed V:

$$\rho = \frac{2gp}{\varsigma V^2}\sqrt{\frac{\varepsilon}{C_{Do}}}\,,\quad p = \frac{m}{S}\,,\quad h = b_1 \ln\frac{a_1}{\rho}\quad , \quad \text{(A4.47)}$$

where $p = m/S$ is the load on a square meter of wing. For a diapason of h = 0–11 km the coefficients a_1 = 1.225, b_1 = 9086.

Results of this computation are presented in Fig. A4.4.

b) Aircraft with turbo-jet engine. The thrust for this engine is

$$T = T_0\frac{\rho}{\rho_0}\,,\quad T'_h = -\frac{T}{b_1}\,,\quad T'_V = 0\,. \quad \text{(A4.48)}$$

Substitute (A4.48) in (A4.33). We obtain

$$V\left(-\frac{T}{b_1} - D'_h\right) = -gD'_V \quad or \quad T = \frac{b_1}{V}\left(gD'_V - VD'_h\right), \quad \text{(A4.48)'}$$

and substituting (A4.44) and (A4.48) in (A4.33), we obtain

$$\frac{1}{p}\left(\frac{V^2}{2b_1}+g\right)\left[C_{Do}-\varepsilon\left(\frac{2pg}{\varsigma\rho V^2}\right)^2\right]=\frac{\overline{T}_0}{b_1\rho_0},\quad where\quad \overline{T}_0=\frac{T_0}{m}\ . \tag{A4.49}$$

We can then find ρ, h from (A4.49)

$$\rho=\frac{2pg\sqrt{\varepsilon}}{\varsigma V^2\sqrt{A_2}},\quad where\quad A_2=C_{Do}-\frac{2p\overline{T}_0}{\rho_0(V^2+2b_1g)}\quad \overline{T}_0=\frac{T_0}{m}\ ,\quad h=b_1\ln\frac{a_1}{\rho}\,. \tag{A4.50}$$

Results of computation for the different p, T = 0.8 *N/kg*, a_1 = 1.225, b_1 = 9086 are presented in Fig. A4.5.

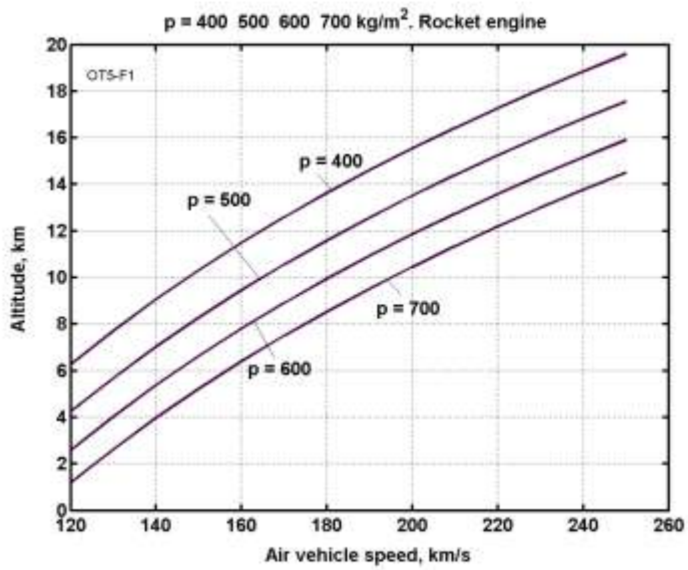


Fig. A4.4. Air vehicle altitude versus speed for wing load p = 400, 500, 600, 700 kg/m^2 and a rocket engine.

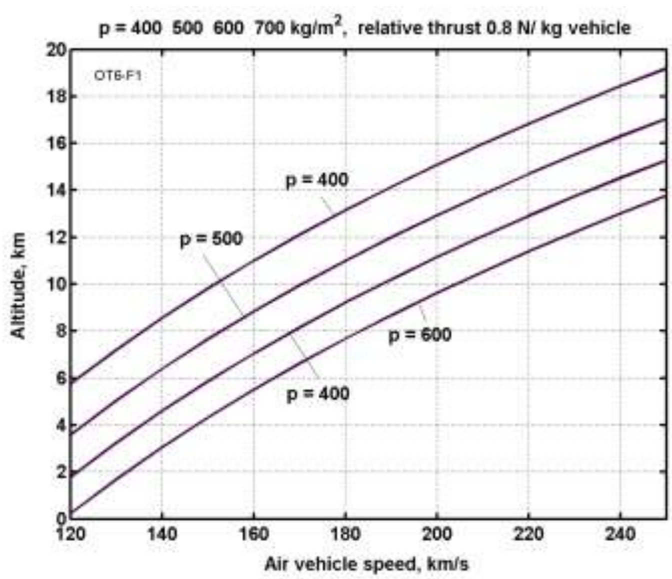


Fig. A4.5. Air vehicle altitude versus speed for wing load *p* = 400, 500, 600, 700 kg/m^2, turbo-jet engine, and relative thrust 0.8 N/kg vehicle.

c) Piston and turbo engines with propeller. All current propeller engines have propellers with variable pitch. The propeller coefficient efficiency, η, approximately is constant. The thrust of this engine is

$$T = \frac{N_0}{V}\frac{\rho}{\rho_0}, \quad T'_V = -\frac{T}{V}, \quad T'_h = -\frac{T}{b_1}, \qquad \text{(A4.51)}$$

where $N_0 = N_e \eta$, N_e is engine power at $h = 0$.

Substituting (A4.44) in (A4.33). We obtain the equation for thrust

$$V\left(\frac{T}{b_1} + D'_h\right) = g\left(\frac{T}{V} + D'_V\right) \quad or \quad T = \frac{b_1 V(gD'_V - VD'_h)}{V^2 - gb_1}. \qquad \text{(A4.51)'}$$

Substitute (A4.44) and (A4.51) in (A4.33). We obtain

$$\frac{V}{p}\left(\frac{V^2}{b_1} - g\right)\left[C_{Do} - \varepsilon\left(\frac{2pg}{\rho V^2}\right)^2\right] = \frac{\overline{N}_0}{\rho_0}\left(\frac{g}{V^2} - \frac{1}{b_1}\right), \quad where \quad \overline{N}_0 = \frac{N_0}{m}, \quad p = \frac{m}{S}. \qquad \text{(A4.52)}$$

We can then find ρ, *h* from (A4.52)

$$\rho = \frac{2pg\sqrt{\varepsilon}}{\varsigma V^2\sqrt{A_3}}, \quad where \quad A_3 = C_{Do} + \frac{p\overline{N}_0}{\rho_0 V^3}, \quad h = b_1 \ln\frac{a_1}{\rho}. \qquad \text{(A4.53)}$$

Results of computation for C_{Do} = 0.025, λ = 10, for different values of *p*, *N* are presented in Fig. A4.6.

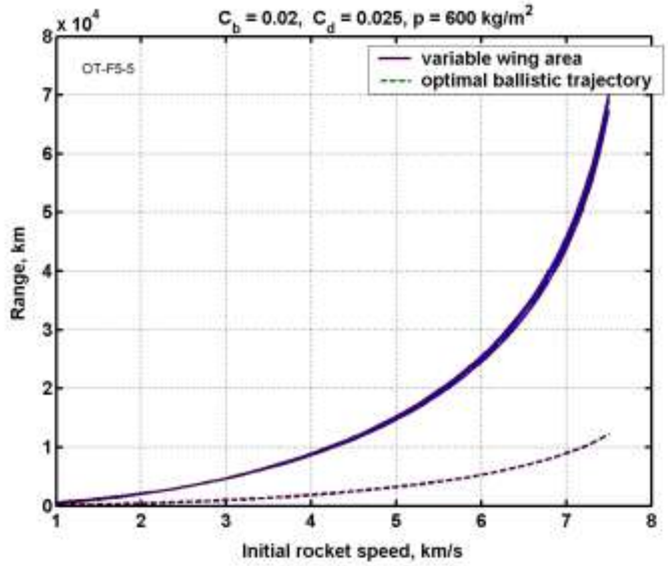


Fig. A4.6. Air vehicle range versus speed for wing load *p* = 250, 300, 350, 400 kg/m^2, piston (propeller) engine, and relative engine power 100 W/kg vehicle.

Hypersonic speed (1 km/s < *V* < 7 km/s).

The lift and drag forces in hypersonic flight are approximately (see (A4.22)')

$$L(\alpha,V,h)=mg-\frac{mV^2}{R}=\varsigma\alpha\frac{a\rho V}{2}S\,,\quad \overline{D}=(C_{DW}+\varepsilon\alpha^2)\frac{a\rho V}{2}S+C_{Db}\frac{a\rho V}{2}S_b\,,$$

$$\alpha=\frac{2p(g-V^2/R)}{\varsigma\rho aV},\quad D=\left[C_{DW}\frac{a\rho V}{2}+\frac{2\varepsilon}{\rho aV}\left(\frac{m(g-V^2/R)}{\varsigma S}\right)^2\right]S+C_{Db}\frac{a\rho V}{2}S_{b,} \qquad \text{(A4.54)}$$

$$or\quad \frac{D}{m}=\left[C_{DW}\frac{q}{p}+\frac{\varepsilon p}{q}\left(\frac{g-V^2/R}{\varsigma}\right)^2\right]+C_{Db}\frac{q}{p_b},\quad q=\frac{\rho aV}{2}\quad .$$

Note

$$D_{0W}=C_{DW}\frac{\rho aV}{2}S\,,\quad D_{0b}=C_{Db}\frac{\rho aV}{2}S_b\,,\quad C_{DW}=4c\,,\quad C_{Db}=2c_b\,,\quad \rho=a_2e^{-\frac{h-11000}{b_2}}\,, \qquad \text{(A4.55)}$$

The derivatives of *D* by *V, h* are

$$D'_V = \frac{D_{0W}}{V} + \frac{D_{0b}}{V} - \frac{2\varepsilon mp}{\varsigma^2 \rho a}\left(g - \frac{V^2}{R}\right)\left(\frac{3}{R} + \frac{g}{V^2}\right),$$

$$D'_h = D'_\rho \rho'_h = -\frac{1}{b_2}\left(D_{0W} + D_{0b} - \frac{2\varepsilon mp(g - V^2/R)^2}{\varsigma^2 \rho a V}\right) \quad . \qquad \text{(A4.56)}$$

a) Rocket engine or hypersonic glider. The derivatives from *T = const* and *T* = 0 are

$$T'_V = 0, \quad T'_h = 0.$$

(A4.57)

Substituting (A4.55) in (A4.56), and expressions (A4.56) and (A4.57) in (A4.37) to find ρ, *h*, we obtain for h > 11,000 m

$$\rho = \frac{2p\sqrt{\varepsilon}}{\varsigma a}\sqrt{A_4}, \quad A_4 = \frac{\left(g - \frac{V^2}{R}\right)\left[g\left(\frac{3}{R} + \frac{g}{V^2}\right) + \frac{1}{b_2}\left(g - \frac{V^2}{R}\right)\right]}{\left(\frac{V^2}{b_2} + g\right)\left[C_{DW} + C_{Db}\left(\frac{S_b}{S}\right)\right]}, \quad h = 11000 + b_2 \ln\frac{a_2}{\rho}, \qquad \text{(A4.58)}$$

where a_2 = 0.365, b_2 = 6997 are coefficients of the exponent atmosphere for the stratosphere at 11 to 60 km.

If we ignore the small term $g\left(\frac{3}{R} + \frac{g}{V^2}\right)$ for *M* > 3 in (A4.58), the equations take the form

$$\rho = \frac{2p(g - V^2/R)\sqrt{\varepsilon}}{\varsigma a}\sqrt{A_5}, \quad A_5 = \frac{1}{C_{Do}(V^2 + gb_2)}, \quad where \quad C_{Do} = C_{0W} + C_{Db}\left(\frac{S_b}{S}\right),$$

where $C_{DW} \approx 4c$. If we ignore the term gb_2 (for *M* > 3), then

$$\rho = \frac{2p(g - V^2/R)}{\varsigma a V}\sqrt{\frac{\varepsilon}{C_{Do}}} \quad . \qquad \text{(A4.59)}$$

In the limit as $R \to \infty$ in (2-54), we find

$$\rho = \frac{2pg}{\varsigma a V}\sqrt{\frac{\varepsilon}{C_{Do}}} \quad . \qquad \text{(A4.59)'}$$

Here $\sqrt{C_{Do}/\varepsilon} = \alpha_{opt}$ is an optimal (maximum C_L/C_D) wing attack angle of the horizontal flight.

Results of the computation in (A4.58) are presented in Fig. A4.7.

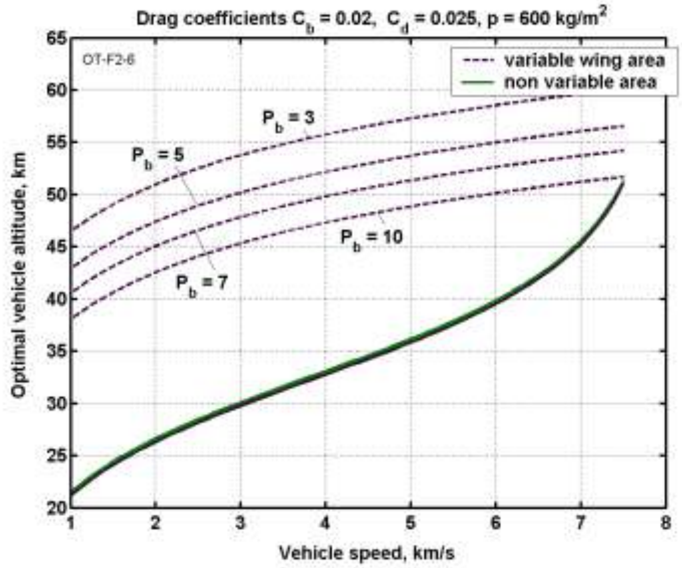


Fig. A4.7. Optimal vehicle altitude versus speed for specific body load P_b = 3, 5, 7, 10 ton/m^2, body drag coefficient C_b = 0.02, wing drag coefficient C_d = 0.025, wing load p = 600 kg/m^2.

b) Ramjet engine. The thrust of the jet engine is approximately ($M < 4$)

$$T = \xi \frac{\rho}{\rho_2} V^2, \quad T'_V = \frac{2T}{V}, \quad T'_h = -\frac{T}{b_2}\ , \qquad \text{(A4.60)}$$

where ξ is a numerical coefficient, ρ_2 is the air density at the lower end of the selected atmospheric diapason (in our case 11 km).

Substituting (A4.60) and (A4.56) in our main equation (A4.33), by repeat reasoning we can obtain the equation for the given engine

$$\rho = \frac{2p\sqrt{\varepsilon}}{\varsigma a}\sqrt{A_6}\ , \quad A_6 = \frac{\left(g - V^2/g\right)\left[g\left(\frac{3}{R}+\frac{g}{V^2}\right)+\frac{1}{b_2}\left(g - V^2/R\right)\right]}{\left[C_{DW}+C_{Db}\left(\frac{S_b}{S}\right)\right]\left[\left(\frac{V^2}{b_2}+g\right)-\frac{2\bar{T}_0 p}{a\rho_0}\left(\frac{V}{b_2}-\frac{2g}{V}\right)\right]}, \quad \bar{T} = \frac{T_0}{m}, \qquad \text{(A4.61)}$$

where T_0 is taken at the lower end of the exponent atmospheric diapason (in our case 11 km). The curve of air density versus altitude h is computed similarly to (A4.58).

Optimal wing area

The lift force and drag of any wing may be written as

$$Y = mg = Y(\alpha, q, S), \quad D = D(\alpha^2, q, S)\,. \qquad \text{(A4.62)}$$

Substituting (A4.62) in (A4.28) and finding the minimum *H* versus *S*, we obtain the equation

$$D + \overline{D}'_{\alpha}\alpha'_{S}S = 0, \quad or \quad D + D'_{S}S = 0, \tag{A4.63}$$

where α is the value found from the first equation (A4.62). Equation (A4.63) is the general equation for the optimal wing area and optimal specific load $p = m/S$ on a wing area.

a) Subsonic speed. Lift force and drag of the subsonic wing are

$$Y = mg = \varsigma\alpha qS \quad or \quad \alpha = \frac{mg}{\varsigma qS}, \quad \overline{D} = \left(C_{Do} + \varepsilon\alpha^2\right)qS \quad D = C_{DW}qS + \varepsilon\left(\frac{mg}{\varsigma}\right)^2\frac{1}{qS}\,, \tag{A4.62)’}$$

where $q = \rho V^2/2$ is a dynamic air pressure for subsonic speed.

Substituting the last equation in (A4.62) into the first equation in (A4.63), we obtain the optimal specific load on the wing area

$$p_{opt} = \frac{\varsigma q}{g}\sqrt{\frac{C_{DW}}{\varepsilon}}\ . \tag{A4.63)’}$$

Substituting α from (A4.62)’ into the last equation in (A4.62)’ and dividing both sides by vehicle mass *m,* we obtain

$$\frac{D}{m} = \left[C_{DW}\frac{1}{p} + \varepsilon\left(\frac{g}{\varsigma q}\right)^2 p\right]q\,. \tag{A4.64}$$

Here *D/m* is specific drag (drag per unit weight for the vehicle). Substituting (A4.63)’ into (A4.64). We abtain the minimum drag for a variable wing

$$\min\left(\frac{D}{m}\right) = 2\frac{g}{\varsigma}\sqrt{\varepsilon C_{DW}}\ , \tag{A4.64)’}$$

where the term on the right is wing drag for the lift of one unit of weight for the vehicle. We discover the important fact than the optimal wing drag of a variable wing does not depend on air speed, it depends only on the geometry of the wing. This may look wrong, but consider the following example. Wing drag is $D = mg/K$, where $K = C_L/C_D$ is the wing efficiency coefficient. The value *D/m* does not depend on speed.

If the air vehicle has a body, the minimum drag is

$$\min\left(\frac{D}{m}\right) = 2\frac{g}{\varsigma}\sqrt{\varepsilon C_{DW}} + C_{Db}\frac{q}{p_b}, \quad q = \frac{\rho V^2}{2}\,. \tag{A4.65}$$

Full vehicle drag depends on speed because the body drag depends on *V*.

Substituting the (A4.63)' term for α into (A4.62)', we obtain the optimal attack angle

$$\alpha_{opt} = \sqrt{\frac{C_{DW}}{\varepsilon}} \,. \qquad \text{(A4.66)}$$

This is the angle of optimal efficiency, but C_{DW} is the wing drag coefficient only when $\alpha = 0$ (not the full vehicle as in conventional aerodynamics). The coefficient of flight efficiency

$$k = \frac{g}{D/m} \quad or \quad k_{\max} = \frac{g}{\min(D/m)} \,. \qquad \text{(A4.67)}$$

b) Hypersonic speed. The equations of wing lift force and wing air drag for hypersonic speed are as follows:

$$Y = \varsigma \alpha q S = m\left(g - \frac{V^2}{R}\right), \quad or \quad \alpha = \frac{p\left(g - V^2/R\right)}{\varsigma q}, \quad \overline{D} = \left(C_{DW} + \varepsilon \alpha^2\right) q S, \quad q = \frac{\rho a V}{2} \,. \qquad \text{(A4.68)}$$

Substituting α from (A4.68) into $\overline{D}$, we obtain

$$D = \left[C_{DW} + \varepsilon \left(\frac{m\left(g - V^2/R\right)}{\varsigma q S} \right)^2 \right] qS \,. \qquad \text{(A4.68)'}$$

Substituting the wing load *p = m/S* into (A4.68)', we obtain

$$\frac{D}{m} = \left[C_{DW} \frac{1}{p} + \varepsilon \left(\frac{g - V^2/R}{\xi q} \right)^2 p \right] q \,. \qquad \text{(A4.69)}$$

To find the minimum the air drag *D* for *p*, we take the derivatives and set them equal to zero, then we obtain

$$p_{opt} = \frac{\varsigma q}{(g - V^2/R)} \sqrt{\frac{C_{DW}}{\varepsilon}} \,. \qquad \text{(A4.70)}$$

Substituting (A4.70) into (A4.69), we find the minimum wing drag

$$\min\left(\frac{D}{m}\right)_w = \frac{2}{\varsigma}\left(g - \frac{V^2}{R}\right)\sqrt{\varepsilon C_{DW}} \,.$$

The sum of the minimum vehicle drag plus body drag is

$$\min\left(\frac{D}{m}\right)=\frac{2}{\varsigma}\left(g-\frac{V^2}{R}\right)\sqrt{\varepsilon C_{DW}}+C_{Db}\frac{q}{p_b}, \quad q=\frac{\rho a V}{2}, \quad p_b=\frac{m}{S_b}. \qquad \text{(A4.71)}$$

Substituting (A4.70) into the term for α in (A4.65), we find the optimal attack angle of a vehicle without a body

$$\alpha_{opt}=\sqrt{C_{DW}/\varepsilon}. \qquad \text{(A4.72)}$$

The coefficient of flight efficiency *k = Y/D* is

$$k=\frac{g-V^2/R}{D/m}, \quad k_{\max}=\frac{g-V^2/R}{\min(D/m)}.$$

For hypersonic speed the coefficients are approximately

$$\varsigma=4, \quad \varepsilon=2, \quad C_{DW}=4c^2, \quad C_{Db}=2c_1^2, \quad C_L=\varsigma\alpha, \quad C_{Do}=C_{DW}+C_{Db}. \qquad \text{(A4.73)}$$

In numerical computation the angle θ can be found from (A4.25) as $\theta = \Delta h/\Delta R_g$.

For the rocket engine or gliding flight we find the following relation: when *S* is optimum (variable), the partial derivatives from (A4.71) are

$$D'_V=-\frac{4V}{\varsigma R}\sqrt{\varepsilon C_{DW}}+C_{Db}\frac{\rho a}{2p_b}, \quad D'_h=-\frac{C_{Db}\rho a V}{2b_2 p_b}.$$

Substituting these into (A4.37), we find the relationship between speed, altitude, and optimal wing load for a hypersonic vehicle with a rocket engine and variable optimal wing:

$$\rho=\frac{8gp_bV\sqrt{\varepsilon C_{DW}}}{\varsigma a C_{Db}R(g+V^2/b_2)}, \quad h=11000+b_2\ln\frac{a_2}{\rho}. \qquad \text{(A4.74)}$$

For $\zeta = 4$, $\varepsilon = 2$ equation (A4.73)’ has the form

$$\rho=\frac{2gp_bV\sqrt{2C_{DW}}}{C_{Db}aR(g+V^2/b_2)}, \quad h=11000+b_2\ln\frac{a_2}{\rho}, \qquad \text{(A4.74)'}$$

Results of computation using (A4.74)’ for $\zeta = 4$, $\varepsilon = 2$, $a_2 = 0.365$, $b_2 = 6997$ and different p_b are presented in Fig. A4.7 (dashed lines). As you see, the variable area wing saves kinetic energy, because its curve is located over an invariable (fixed) wing. This is advantageous only at orbital speed (7.9 km/s) because no lift force is necessary.

Estimation of flight range

Air and space vehicles without thrust

The aircraft range can be found from equation (A4.26)

$$R_a = \int_{V_1}^{V_2} \frac{mVdV}{T - D - mg\theta}, \quad V_1 > V_2 \quad or \quad R_a = \int_{V_2}^{V_1} \frac{VdV}{D/m + g\theta}, \quad if \quad T = 0. \qquad \text{(A4.75)}$$

Consider a missile with the *optimal variable wing* in a descent trajectory with thrust $T = 0$.

a) Make the simplest estimation using equations for kinetic energy from classical mechanics. Separate the flight into two stages: hypersonic and subsonic. If we have the ratio of vehicle efficiency $k_1 = C_L / C_D$, $k_2 = C_L / C_D$, where k_1, k_2 are the ratios of flight efficiency for the hypersonic and subsonic stages respectively, we find the following equations for a range in each region:

$$\frac{m}{2}\left(V_1^2 - V_2^2\right) = \frac{m(g - V^2/R)}{k_1} R_1, \quad R_1 = \frac{k_1\left(V_1^2 - V_2^2\right)}{2(g - V^2/R)}, \quad R_2 = k_2 h, \quad R_a = R_1 + R_2,$$

Or more exactly

$$d\left(\frac{mV^2}{2}\right) = \frac{m(g - V^2/R)}{k_1} dR_1, \quad R_1 = -\frac{k_1 R}{2} \ln\left(\frac{g - V_2^2/R}{g - V_1^2/R}\right), \qquad \text{(A4.76)}$$

where R_1 is the hypersonic part of the range, R_2 is the subsonic part of the range, V_1 is the initial (maximum) vehicle hypersonic speed, V_2 is a final hypersonic speed, and h is the altitude at the initial stage of the subsonic part of the trajectory.

b) To be more precise. Assume in (A4.75) $\rho = const$ (taking average air density).

1. For the hypersonic part of the trajectory: substitute (A4.71) into (A4.76). We then have

$$R_{1H} = \int_{V_1}^{V_2} \frac{VdV}{aV^2 + bV + c}, \quad or \quad R_{1H} = \int_{V_1}^{V_2} \frac{VdV}{X}, \quad where \quad X = aV^2 + bV + c,$$
$$a = \frac{2\sqrt{\varepsilon C_{DW}}}{\xi R}, \quad b = -C_{Db}\frac{\rho a}{2p_b}, \quad c = \frac{T}{m} - \frac{2g}{\varsigma} - g\theta, \quad \Delta = 4ac - b^2, \qquad . \qquad \text{(A4.77)}$$
$$R_{1H} = \left[\frac{1}{2a}\ln X - \frac{b}{2a}\int \frac{dV}{X}\right]_{V_1}^{V_2}, \quad \int \frac{dV}{X} = \frac{2}{\sqrt{\Delta}} \arg\tan \frac{2aV + b}{\sqrt{\Delta}} \quad for \quad \Delta \ge 0,$$
$$\int \frac{dV}{X} = -\frac{2}{\sqrt{-\Delta}} \arg\tanh \frac{2aV + b}{\sqrt{-\Delta}} = \frac{1}{\sqrt{-\Delta}} \ln \frac{2aV + b - \sqrt{-\Delta}}{2aV + b + \sqrt{-\Delta}} \quad for \quad \Delta \le 0.$$

2. For the subsonic part of the trajectory: substitute (A4.65) into (A4.75). We then have

$$R_{1S} = -\frac{1}{2C_2}\ln\left|\frac{C_1 - C_2V_2^2}{C_1 - C_2V_1^2}\right| \ , \tag{A4.78}$$

where the values for C_1, C_2 are

$$C_1 = \frac{T}{m} - g\left(\frac{2\sqrt{\varepsilon C_{DW}}}{a\varsigma} + \theta\right), \quad C_2 = C_{Db}\frac{\rho}{2p_b} \ . \tag{A4.79}$$

The trajectory (without the rocket part of the trajectory) is

$$R_1 = R_{1H} + R_{1S} \quad or \quad R_g = R_{1H} + R_{1S} + R_2 \, . \tag{A4.80}$$

where $R_2 = k_2h$ computed for altitude h at the end of the kinetic part of the subsonic trajectory.

3. The ballistic trajectory of a wingless missile without atmosphere drag is

$$h = \frac{gt^2}{2}, \quad t = \sqrt{\frac{2h}{g}}, \quad R_b = V_1t = V_1\sqrt{\frac{2h}{g}}, \quad V_i = \sqrt{V_1^2 + V_y^2}, \quad V_y^2 = 2h(g - V^2/R)\,, \tag{A4.81}$$

where h is the initial altitude, V_1 is the initial horizontal speed of the wingless missile at altitude h, V_y is initial (shot) vertical speed at $h = 0$, V_i is the full initial (shot) speed at $h = 0$.

For the hypersonic interval 5 < V < 7.5 km/s, we can use the more exact equation

$$R_b = V_1\sqrt{\frac{2h}{(g - V_1^2/R)}} \ , \tag{A4.82}$$

where R = 6378 km is the radius of Earth. The full range of a ballistic rocket plus the range of a winged missile is

$$R_f = R_b + R_a + R_g, \tag{A4.83}$$

where $R_g = kh$ is the vehicles gliding range from the final altitude h_2 (see Fig. A4.11) with aerodynamic efficiency k.

The classical method finding of the optimal shot ballistic range for spherical Earth without atmosphere is

$$R_b = 2R\beta_{opt}\,, \quad \tan\beta_{opt} = \frac{\nu_A}{2\sqrt{1-\nu_A}}, \quad \nu_A = \frac{V_A^2}{V_c^2} \ , \tag{A4.84}$$

where β_{opt} is the optimal shot angle, V_A is the shot projectile speed, and V_c is an orbital speed for a circular orbit at a given altitude.

4. **Cannon projectile**. We divide the distance into three sub-distances: 1) 1.2M < M, 2) 0.9M < M < 1.2M, 3) 0 < M < 0.9M. The range of the wing cannon projectile may be estimated using the equation

$$R=\frac{k_1}{2g}(V_1^2-V_2^2)+\frac{k_2}{2g}(V_2^2-V_3^2)+\frac{k_3}{2g}(V_3^2-V_0^2),\quad where\quad 0<V_0<V_3<V_2<V_1\ , \qquad \text{(A4.85)}$$

where k_1, k_2, k_3 are the average aerodynamic efficiencies for sub-distances 1, 2, 3 respectively. Conventionally, these coefficients have the following values: subsonic k_3 = 8–15, near sonic k_2 = 2–3, supersonic and hypersonic k_1 = 4–9. If $V > 600$ m/s, the first term in (A4.85) has the greatest value and we can use the more simple equation for range estimation:

$$R=\frac{k_1}{2g}V_1^2. \qquad \text{(A4.84)'}$$

At the top of its trajectory, a modern projectile can have an additional impulse from small rocket engines. Their weight is 10–15% of the full mass of the projectile and increases the maximum range by 7–14 km. In this case we must substitute $V = V_1 + dV$ into (A4.84)', where dV is the additional impulse (150–270 m/s).

Subsonic aircraft with thrust. Horizontal flight

The optimal climb and descent of a subsonic aircraft with a constant mass and fixed wing is described by equations (A4.50) and (A4.47). Any given point in a climb curve may be used for horizontal flight (with different efficiency). We consider in more detail the horizontal flight when the aircraft mass decreases because the fuel is spent. This consumption may reach 40% of the initial aircraft mass. The optimal horizontal flight range may be computed in the following way:

$$dR=Vdt,\quad dt=\frac{dm}{c_sT}=\frac{gdm}{c_sD},\quad dR=\frac{gV}{c_sD(m)}dm,\quad R=\frac{gV}{c_s}\int_{m_k}^{m}\frac{dm}{D(m)}, \qquad \text{(A4.86)}$$

where m is fuel mass, c_s is fuel consumption, kg/s/ kg thrust.

a) For a **fixed wing**, we have (from (A4.44))

$$D=C_{Do}qS+\frac{\varepsilon}{qS}\left(\frac{g}{\varsigma}\right)^2m^2,\quad where\quad C_{Do}=C_{Dw}+C_{Db}\left(\frac{S_b}{S}\right),\quad q=\frac{\rho V^2}{2} \qquad \text{(A.87)}$$

Substituting (A4.87) into (A4.86), we obtain

$$R=\frac{gV}{c_s\sqrt{C_1C_2}}\arg\tan\frac{\sqrt{C_1/C_2}\,(m-m_k)}{1+(C_1/C_2)mm_k},\quad where\quad C_1=\frac{\varepsilon}{qS}\left(\frac{g}{\varsigma}\right)^2,\quad C_2=C_{Do}qS. \qquad \text{(A4.88)}$$

b) For a **variable wing** we have (from (A4.65)

$$R=\frac{gV}{c_sC_1}\ln\frac{C_1m-C_2}{C_1m_k-C_2},\quad where\quad C_1=2\frac{g}{\varsigma}\sqrt{\varepsilon C_{DW}},\quad C_2=C_{Db}qS_b\,,\quad \rho=\rho_0e^{-h/b_1}.$$

(A4.89)

Results of the computation are presented in Fig. A4.8. The aircraft have the following parameters: C_{DW} = 0.02; C_{Db} = 0.08; b_1 = 9086; S = 120 m^2; m = 100 tons, m_k = 80 tons, c_s = 0.00019 kg/s/kg thrust; wing ratio λ = 10.

As you see, the specific fuel consumption does not depend on speed and altitude, a good aircraft design reaches the maximum range only at one point, in one flight regime: when the aircraft flies at the maximum speed possible for the critical Mach number, at the maximum altitude possible for that engine. The deviation from this point decreases in the range in 5–10–15 percent or more. The variable wing increases efficiency of the other regime, which that approximately reduces the losses by a half.

The coefficient of flight efficiency may be computed using equation $k = g/(D/m)$, where the values

$$\frac{D}{m} = C_{DW}\frac{q}{p} + \frac{\varepsilon p}{q}\left(\frac{g}{\varsigma}\right)^2 + C_{Db}\frac{q}{p_b}, \quad \left(\frac{D}{m}\right)_1 = 2\frac{g}{\varsigma}\sqrt{\varepsilon C_{DW}} + C_{Db}\frac{q}{p_b}, \qquad \text{(A4.90)}$$

apply for fixed and variable wings respectively. Results of computation are presented in Fig. A4.9. The curve of the variable wing is the round curve of the fixed wing.

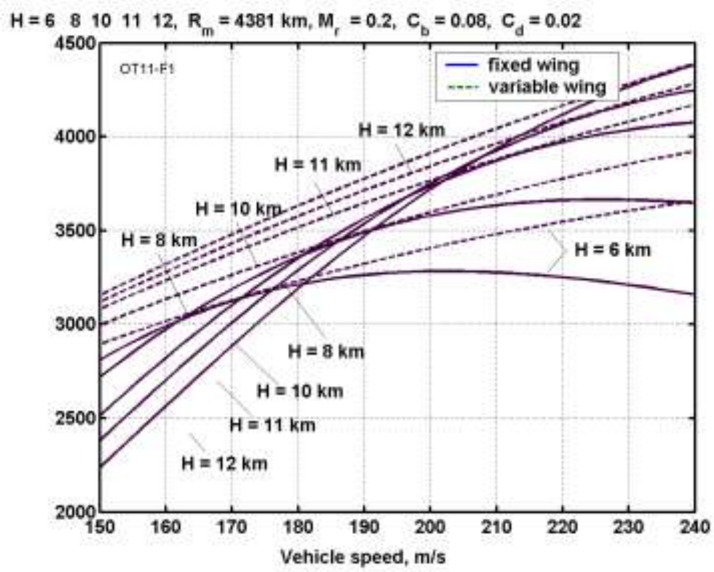


Fig. A4.8. Aircraft range for altitude H = 6, 8, 10, 11, 12 km; maximum range R_m = 4361 km; relative fuel mass M_r = 0.2; body drag coefficient C_b = 0.08; wing drag coefficient C_d = 0.02.

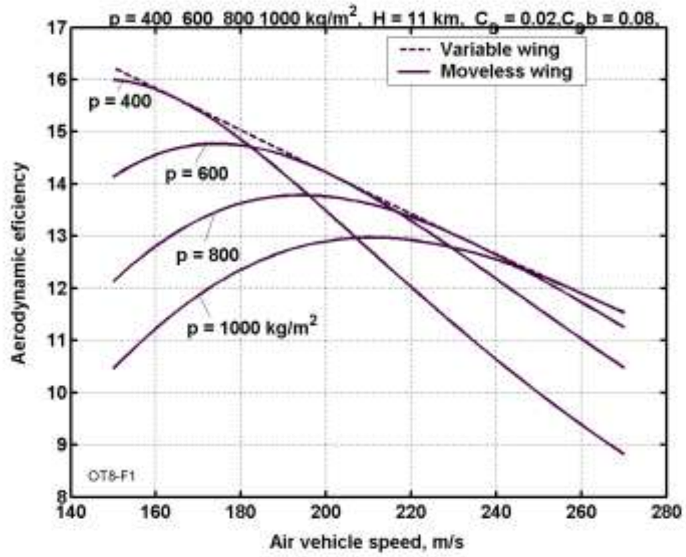


Fig. A4.9. Aerodynamic efficiency of non-variable and variable wings for wing load *p* = 400, 600, 800, 1000 kg/m^2, wing drag C_D = 0.02, body drag C_{Db} = 0.08, wing ratio 10.

Optimal engine control for constant flight pass angle

Let us to consider equations (A4.1) – (A4.5) for a constant angle of trajectory, θ = const. Substituting θ = constant, thrust $T = V_e\beta$, and a new independent variable $s = Vt$ (where s is the length of the trajectory) into the equation system (A4.1) – (A4.5). We obtain the following equations

$$\frac{dL}{ds} = \cos\theta \quad ,$$
$$\frac{dh}{ds} = \sin\theta \quad ,$$
$$\frac{dV}{ds} = \frac{V_e(h,V)\beta - \overline{D}(\alpha,V,h)}{mV} - \frac{g}{V}\sin\theta \,, \qquad \text{(A4.91) – (A4.96)}$$
$$\frac{dm}{ds} = -\frac{1}{V}\beta \quad ,$$
$$Y(\alpha,V,h) - gm\cos\theta + \frac{mV^2}{R} + 2mV\omega_E\cos\varphi_E = 0 \,,$$
$$0 \le \beta \le \beta_{\max} \quad .$$

Equation (A4.95) is used to substitute for α in equation (A4.93) and for a change of air drag

$$\overline{D}(\alpha,V,h) = D(V,h). \qquad \text{(A4.97)}$$

We find a non-linear system with a linear fuel control β. This means the system can have a singular solution.

Solution

Consider the maximum range for vehicles described by equation (A4.91) – (A4.96).

Let us write the Hamiltonian *H*

$$H = \cos\theta + \lambda_1 \sin\theta + \lambda_2\left[\frac{V_e(h,V)\beta - D(V,h)}{mV} - \frac{g}{V}\sin\theta\right] - \lambda_3\frac{1}{V}\beta \quad , \tag{A4.98}$$

where $\lambda_1(s), \lambda_2(s), \lambda_3(s)$ are unknown multipliers. Application of conventional methods gives

$$\dot{\lambda}_2 = -\frac{\partial H}{\partial V} = -\lambda_2\left[\left(-\frac{1}{V^2}\right)\left(\frac{V_e\beta - D(V,h)}{m} - g\sin\theta\right) - \frac{D'_V}{mV}\right] - \lambda_3\frac{1}{V^2}\beta \quad ,$$

$$\dot{\lambda}_3 = -\frac{\partial H}{\partial m} = \lambda_2\frac{V_e\beta - D}{m^2V}, \qquad \text{(A4.99) – (A4.101)}$$

$$\beta = \max_{\beta} H = \beta_{\max} sign\left[\lambda_2 V_e - \lambda_3 m\right] \quad .$$

Where D'_V is the first partial derivate of *D* by *V*.

The last equation shows that the fuel control β can have only two values, $\pm\beta_{\max}$. We consider the singular case when

$$A = \lambda_2 V_e - \lambda_3 m \equiv 0 . \tag{A4.102}$$

This equation has two unknown variables, λ_2 and λ_3, and does not contain information about fuel control β.

The first two equations (A4.91) – (A4.92) do not depend on variabls and can be integrated

$$L = s\cos\theta, \tag{A4.103}$$

$$H = s\sin\theta. \tag{A4.104}$$

In accordance with the References[2] let us differentiate equation (A4.102) by the independent variable *s*. After substitution into equations (A4.93) – (A4.95), (A4.97), (A4.99), (A4.100), (A4.102), and (A4.104) we obtain the relation for $\lambda_2 \neq 0$, $\lambda_3 \neq 0$:

$$\dot{A} = VD - mVD'_m + V_e(-D - mg\sin\theta + VD'_V) - VV'_{e,V}(D - mg\sin\theta) + mV^2V'_{e,s} = 0 . \tag{A4.105}$$

This equation also does not contain β, however it does contain an important relation between variables *m*, *h* and *V*, on an optimal trajectory. This is a 3-dimentional surface. If we know

$$D = D(h,V) , \tag{A4.106}$$

$$V_e = V_e(h,V) , \tag{A4.107}$$

The mass of our apparatus *m,* and its altitude *h*, we can find the optimal flight speed. This means we can calculate the necessary thrust and the fuel consumption for every point *m*, *h*, *V* (Fig. A4.10).

If we want to find an equation for the fuel control β, we continue to differentiate equation (A4.105) to find the independent variable *s* and substitute in equations (A4.91) – (A4.104). If we calculate the relation for β, if $\lambda_2 \neq 0$, $\lambda_3 \neq 0$, V_e = const, then

$$\beta = \frac{\dot{A}'_V (D + mg\sin\theta) - mV\dot{A}'_s}{V_e \dot{A}'_V - m\dot{A}'_m}, \quad \text{(A4.108)}$$

where

$$\dot{A}'_V = \frac{\partial}{\partial V}\left(\frac{dA}{ds}\right), \quad \dot{A}'_s = \frac{\partial}{\partial s}\left(\frac{dA}{ds}\right). \quad \text{(A4.109)}$$

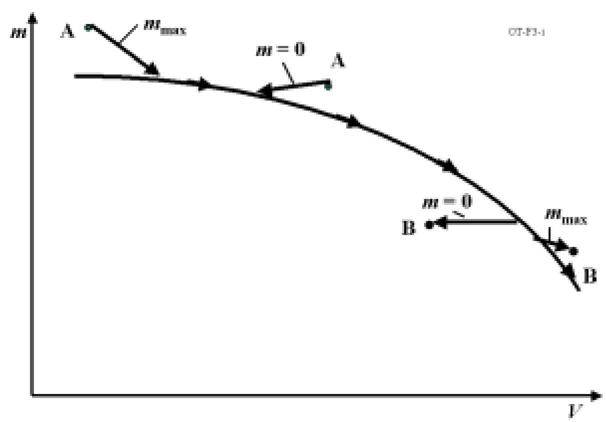


Fig. A4.10. Optimal fuel consumption of flight vehicles.

The necessary condition of the optimal trajectory as it is shown in the References[2–8] (see for example, equation (4.2)[4]) is

$$-(-1)^k \frac{\partial}{\partial\theta}\left[\frac{d^{2k}}{ds^{2k}}\left(\frac{\partial H}{\partial s}\right)\right] \geq 0\,. \quad \text{(A4.110)}$$

where *k* = 1.

If the flight is horizontal ($\theta = 0$), the expression (A4.108) is very simply

$$\beta = \frac{D}{V_e}\ . \quad \text{(A4.111)}$$

This means the thrust equals the drag, a fact that is well known in aerodynamic science.

To obtain the specific equations for different forms of drag and thrust, we must take formulas (or graphs) for subsonic, transonic, supersonic and hypersonic speed for thrust and substitute them into the equations (A4.105) and (A4.108).

Simultaneous optimization of the path angle and fuel consumption

Consider the case where the path angle and the fuel consumption are simultaneously optimized.

In this case the general equations (A4.1) – (A4.5) have the form:

$$\frac{dL}{ds}=1\,,$$
$$\frac{dh}{ds}=\theta\,,$$
$$\frac{dV}{ds}=\frac{V_e(h,V)\beta-D(m,V,h)}{mV}-\frac{g}{V}\theta\,, \qquad \text{(A4.112) – (A4.116)}$$
$$\frac{dm}{ds}=-\frac{1}{V}\beta\,,$$
$$Y(\alpha,V.h)=mg+\frac{mV^2}{R}+2mV\omega_E\cos\varphi_E\,.$$

Let us write the Hamiltonian

$$H=1+\lambda_1\theta+\lambda_2\left(\frac{V_e(h,V)\beta-D(m,V,h)}{mV}-\frac{g}{V}\theta\right)-\lambda_3\frac{1}{V}\beta\,. \qquad \text{(A4.117)}$$

The necessary conditions of optima give

$$A=\frac{\partial H}{\partial\theta}=V\lambda_1-g\lambda_2=0\,,$$
$$B=\frac{\partial H}{\partial\beta}=V_e\lambda_2-m\lambda_3=0\,, \qquad \text{(A4.118) – (A4.119)}$$

The lambda equations are

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial h} = -\lambda_2 \frac{V'_{e,h}\ \beta - D'_h}{mV},$$
$$\dot{\lambda}_2 = -\frac{\partial H}{\partial V} = -\lambda_2 \left[\frac{(V'_{e,V}\beta - D'_V)V - (V_e\beta - D)}{mV^2} + \frac{g}{V^2}\theta\right] - \lambda_3 \frac{1}{V^2}\beta, \quad \text{(A4.120) – (A4.122)}$$
$$\dot{\lambda}_3 = -\frac{\partial H}{\partial m} = \lambda_2 \frac{V_e\beta - D + mD'_m}{m^2 V}.$$

If we differentiate *A* (A4.118), from *dA/ds* = 0, we find the optimal fuel consumption

$$\beta = \frac{gVD'_V - V^2 D'_h}{g(V_e + V'_{e,V}V) - V'_{e,h}V^2}. \qquad \text{(A4.123)}$$

Then we differentiate *B* (A4.119), from *dB/ds* = 0 we find the optimal path angle

$$\theta = \frac{V'_{e,V}D - V_e D'_V - V_e D/V - D + mD'_m}{m(g + V'_{e,h}V - V'_{e,V}g)}. \qquad \text{(A4.124)}$$

We have used the conventional forms for the partial derivatives in (A4.120)–(A4.124) as in the earlier sections of the chapter (see for example (A4.51)).

If we know from analytical formulas or graphical functions *Ve*, *D*, *Y* we can find the optimal trajectory of the air vehicle.

In the general case, this trajectory includes four parts:

1. Moving between limitations θ and β.
2. Moving between one limitation θ or β and one optimal control β or θ.
3. Moving simultaneously with both optimal controls θ and β.
4. Moving at a given point along one limitation and/or both limitations

.

Application to aircraft, rocket missiles, and cannon projectiles

A) Application to rocket vehicles and missiles.

Let us apply the previous results to typical current middle- and long-distance rockets with warheads. We will show: if the warhead has wings and uses the optimal trajectory, the range of the warhead (or its useful load) is increased dramatically in most cases. We will compute the optimal trajectories for a rocket-launched warhead at a particular altitude (20–60 km) and speed (1–7.5 km/s). Point *B* is located on the curve (A4.58) for a fixed wing and on curve (A4.73)' for a variable wing (Fig. A4.11). Further, the winged warhead flies (descends) along the optimal trajectory *BD* (Fig. A4.58) according to equations (A4.58) (fixed wing) or equations (A4.73)' (variable wing) respectively. When the speed is reduced by a small amount (for example, 1 km/s) (point *D* in Fig. A4.11), the winged warhead glides (distance *DE* in Fig. A4.11).

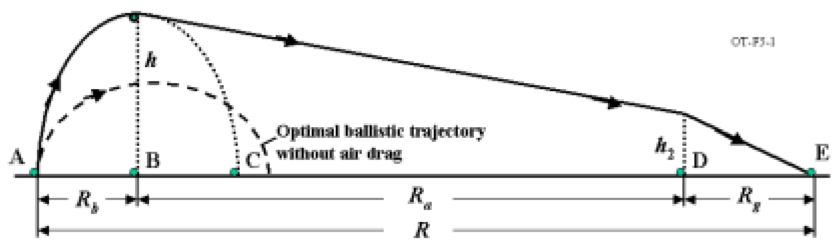


Fig. A4.11. Trajectory of flying vehicles.

The following equations are used for computation:

1. ***The optimal trajectory for a fixed wing space vehicle.***

a) Equation (A4.58) is used to calculate $h = h(V)$ to find the optimal trajectory of a warhead with a non- variable fixed wing in the speed interval 1 < V < 7.5 km/s. The result is presented in Fig. A4.7.

b) Equation (A4.54) gives the magnitude (D/m).

c) The equation (A4.75) in the form

$$\Delta R_a = \frac{V\Delta V}{(D/m)+g\theta}, \quad R_a = \Sigma \Delta R_a, \quad \theta = -\frac{\Delta h}{\Delta R_a}, \quad k = \frac{g - V_0^2/R}{D/m}, \quad R_g = h_0 k, \qquad \text{(A4.125)}$$

is used for computation in the intervals R_a, R_g (Fig. A4.11). Here R_g is the range of a gliding vehicle.

d) Equation (A4.75) is used to calculate R_b in the launch interval *AB* (Fig. A4.11).

e) The full range, *R*, of a warhead with a fixed wing and the full ballistic warhead range, R_w, are

$$R = R_b + R_a + R_g, \quad R_w = 2R_b. \qquad \text{(A4.126)}$$

f) Equation (A4.84) is used to calculate the optimal **ballistic** trajectory of a shot without air drag (a vehicle **without** wings). The range of this trajectory, as it is known, may be significantly more than the range in the atmosphere.

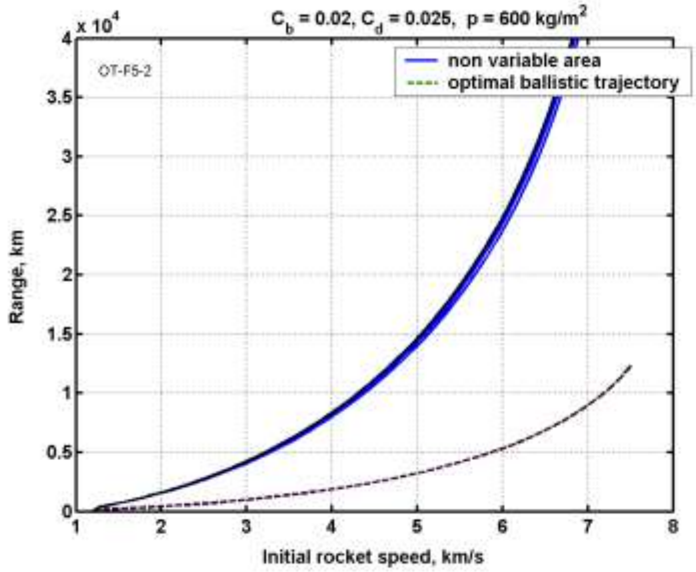


Fig. A4.12. Range of NON-VARIABLE wing vehicle for body drag coefficient $C_b = 0.02$, wing drag coefficient $C_d = 0.025$, wing load $p = 600$ kg/m^2.

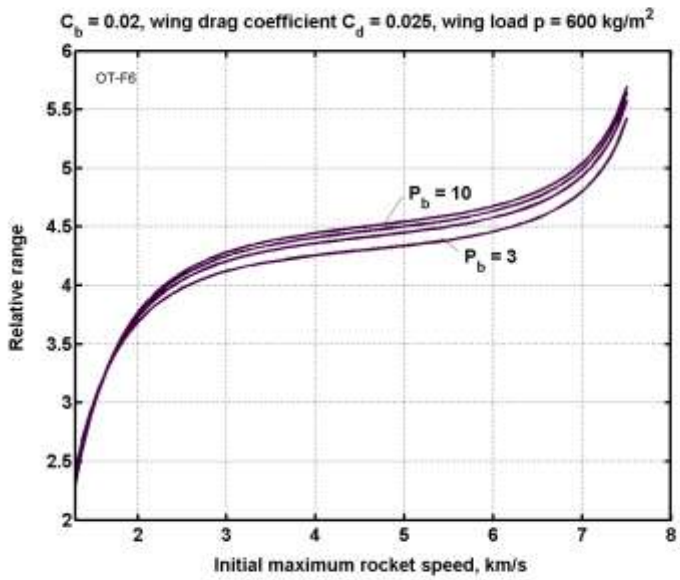


Fig. A4.12. The relative range of a non-variable wing vehicle for the body drag coefficient $C_b = 0.02$, wing drag coefficient $C_d = 0.025$, wing load $p = 600$ kg/m^2, body load $P_b = 3$–10 ton/m^2.

The results are presented in Fig. A4.12. Computation of the relative range (for different p_b) using the formula

$$R_r = \frac{R_f}{R_b} \tag{A4.127}$$

is presented in Fig. A4.12. The optimal range of the winged vehicle is approximately 4.5 times that of the ideal ballistic rocket computed without air drag. In the atmosphere this difference will be significantly more.

2. *Rockets, missiles and space vehicles with* variable *wings*

The computation is the same. For computing ρ, *h*, *D/m* we can use equations (A4.73)' and (A4.71) respectively. The results for different body loads are presented in Fig. A4.7. The optimal trajectories of vehicles with variable wing areas have less slope. This means the vehicle loses less energy when it moves. It travels above the optimal trajectory of a vehicle with fixed wings, which means it needs a lot more time (10–20) and more wing area than a fixed wing space vehicle (Fig. A4.14). The computation of the optimal variable wing area is presented in Fig. A4.15. The relative range (equation (A4.127)) is presented in Fig. A4.16.

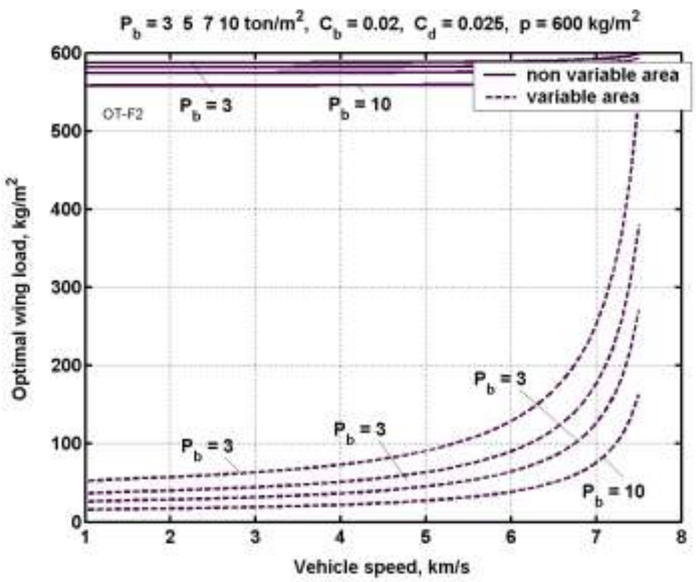


Fig. A4.14. Optimal wing load versus speed for specific body load $P_b = 3, 5, 7, 10$ ton/m^2, body drag coefficient $C_b = 0.02$, wing drag coefficient $C_d = 0.025$, wing load $p = 600$ kg/m^2.

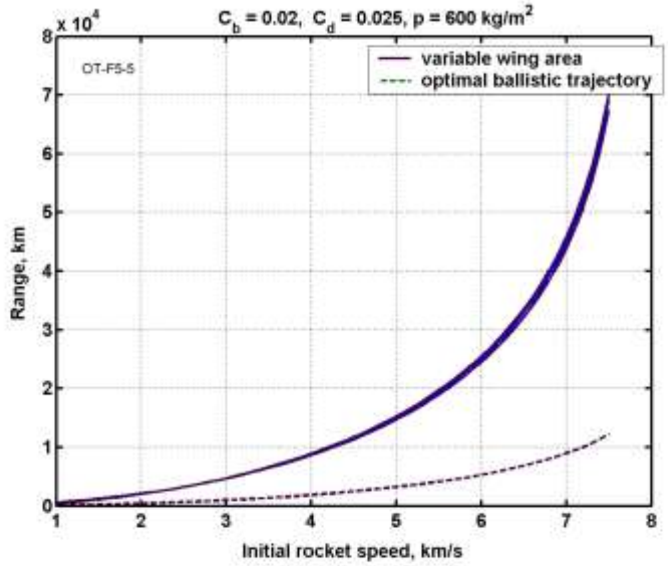


Fig. A4.15. Range of a variable wing vehicle for the body drag coefficient $C_b = 0.02$, the wing drag coefficient $C_d = 0.025$, the wing load $p = 600$ kg/m^2.

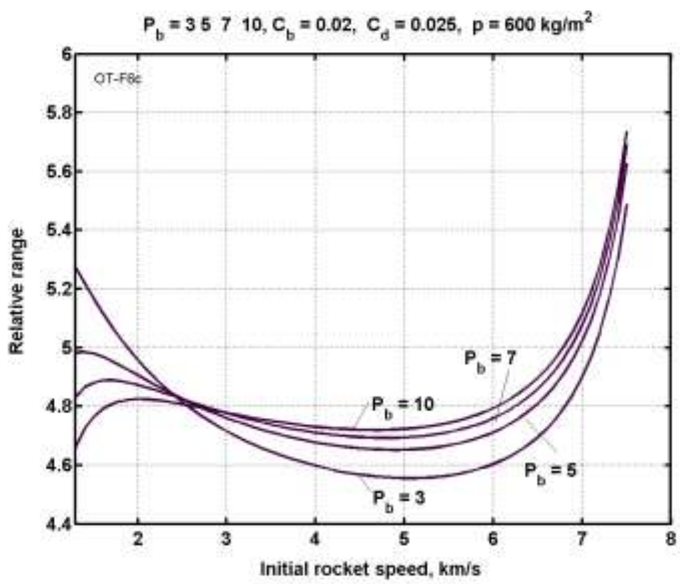


Fig. A4.16. Relative range of variable wing vehicle for the body drag coefficient $C_b = 0.02$, the wing drag coefficient $C_d = 0.025$, the wing load $p = 600$ kg/m^2, the body load $P_b = 3–10$ ton/m^2.

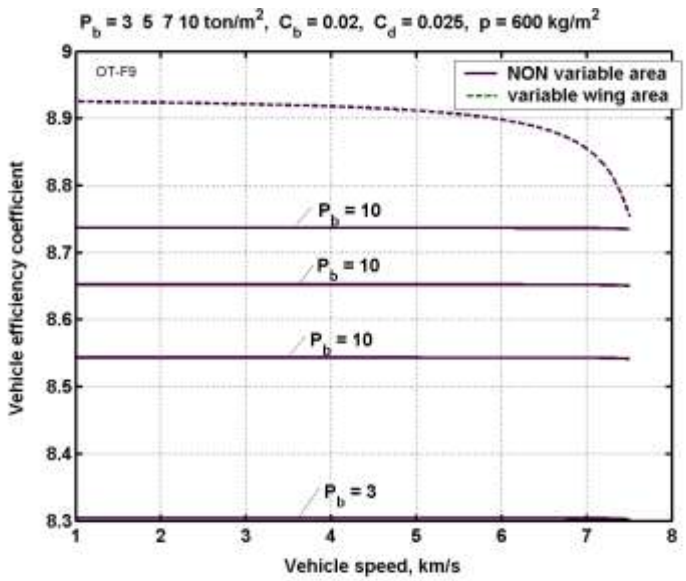


Fig. A4.17. Vehicle efficiency coefficient versus speed for specific body load P_b = 3, 5, 7, 10 ton/m^2, body drag coefficient C_b = 0.02, wing drag coefficient C_d = 0.025, wing load p = 600 kg/m^2.

The aerodynamic efficiency of vehicles with fixed (for different p_b bodies) and optimal variable wings computed using equations (A4.125) and (A4.67) respectively is presented in Fig. A4.12. The difference between vehicles with fixed and variable wings reaches 0.2–0.6 . The slope of the trajectory to horizontal is small (Fig. A4.18).

The range of the **fixed** wing vehicle computed using equation (A4.125) is presented in Fig. A4.12. The range of the **variable** wing vehicle computed using equation (A4.126) is presented in Fig. A4.15. The curve is practically the same (see Figs. A4.12 and A4.15).

3. ***Increasing the rocket payload for the same range.*** If we do not need to increase the range, the winged vehicle can be used to increase the payload, or to save rocket fuel. We can change the mass of the fuel or the payload. The additional payload may be estimated by the following equation

$$\mu = 1 - e^{-\frac{\Delta V}{V_e}}, \qquad \text{(A4.128)}$$

where $\mu = m/m_b$ is relative mass (the ratio of rocket mass of the winged vehicle to the ballistic rocket), $\Delta V = V_b - V$ is the difference between the optimal ballistic rocket speed (equation (A4.84)) and the rocket with a winged vehicle (equation (A4.126)) for given range (see Fig. A4.12). Results of computation are presented in Fig. A4.19. The mass of the rocket with a winged vehicle may be only 20–35% of the optimal ballistic rocket flown without air drag.

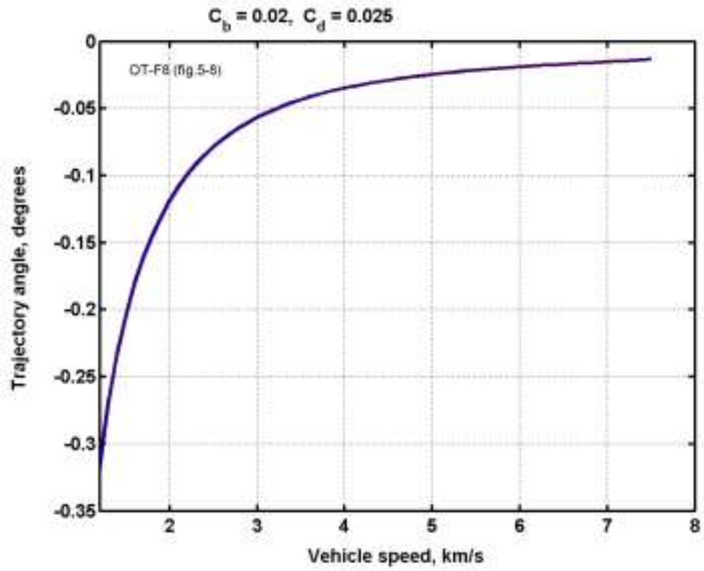


Fig. A4.18. Trajectory angle versus speed for body drag coefficient C_b = 0.02, wing drag coefficient C_d = 0.025.

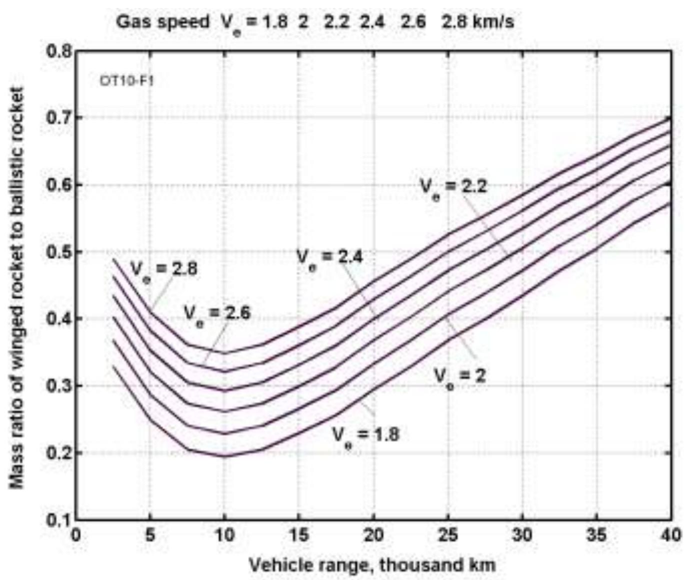


Fig. A4.19. Ratio of mass of winged rocket to ballistic rocket for specific engine run-out gas speed V_e = 1.8, 2, 2.2, 2.4, 2.6 and 2.8 km/s.

***Conclusion*:** The winged air-space vehicle has a range that is greater by a minimum of 4.5–5 times than an optimal shot ballistic space vehicle. The variable wing improves the aerodynamic efficiency by 3–10% and also improves the range. An optimal variable wing requires a large wing area. If you do not need to increase the range, you may instead increase payload.

B) Application to cannon wing projectiles

Properties of a typical current cannons are shown in Table A4.1.

Table A4.1. Properties of current typical Cannons.

Name	caliber, mm	Nozzle speed, m/s	Mass of projectile, kg	Range, km	RAP, km
M107	175	509–912	67	15–33	
SD-203	203	960	110	37.5	
2S19	155	810	43.6	24.7	
2S1	122	690–740	21.6	-	
S-23	180	-	-	30.4	43.8
2A36	152	-	-	17.1	24
D-20	152	600–670	43.5–48.8	20	

Issue: Jane's

The computations using equation (A4.84)' for different *k* and RAP with dV = 270 m/s are presented in Figs. A4.20 and A4.21.

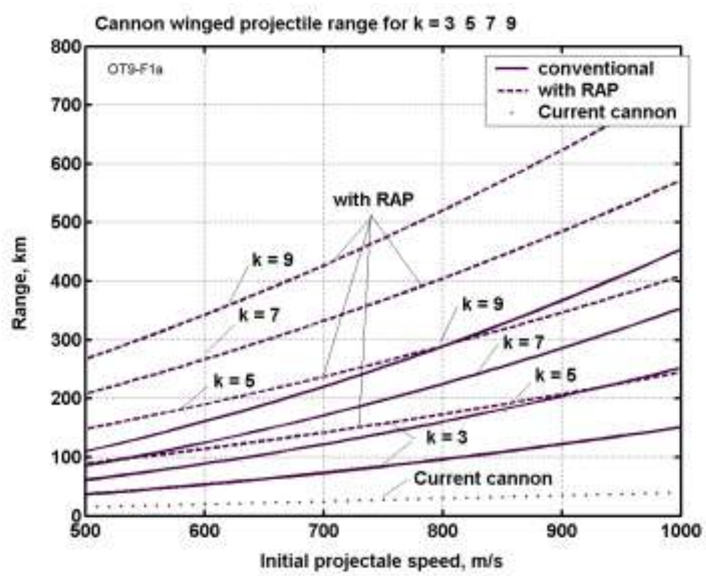


Fig. A4.20. Cannon winged projectile range for average aerodynamic efficiency *k* = 3, 5, 7, 9.

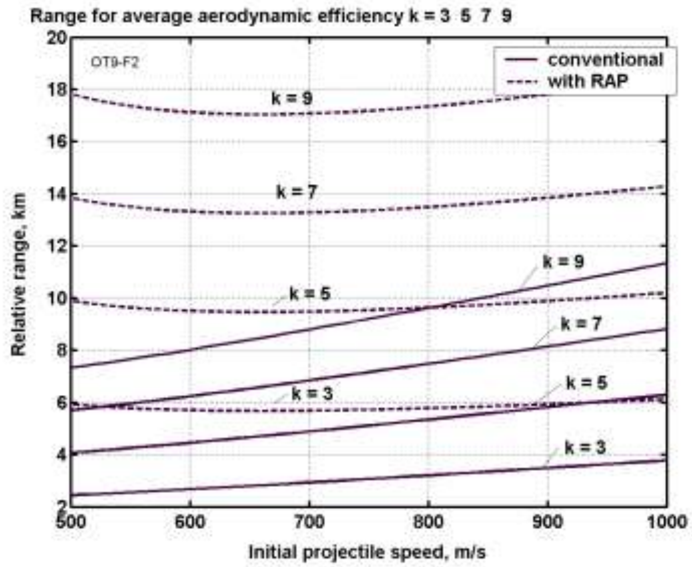


Fig. A4.21. Cannon winged projectile relative range for average aerodynamic efficiency *k* = 3, 5, 7, 9.

Conclusion. As you see (Figs. A4.20, A4.21), the winged projectile increase its range 3–9 times (from 35 up to 360 km, $k = 9$). The projectile with RAP increases its range 5–14 (from 40 up to 620 km, $k = 9$) . Winged shells have another important advantage: they do not need to rotate. We can use a barrel with a smooth internal channel. This allows for an increase in projectile nozzle speed of up to 2 km/s and in shell range of up to 1000 km ($k = 5$).

C) Application to current aircraft.

We can use equations (A4.88) and (A4.89) for computations for typical passenger airplanes (Figs. A4.22, A4.23, A4.24, and A4.8), where all values are divided by the maximum range R_m = 4381 km (for a fuel mass that is 20% of to vehicle mass) at a speed of V = 240 m/s, and altitude H = 12 km. The speed is limited by the critical Mach number ($V < M = 0.82$), and the altitude is limited by the engine trust, when engine stability is such that it works in a cruise regime. Fig. A4.22 shows the typical long-range trajectory of aircraft.

Conclusion: The best flight regime for a given air vehicle (closed to Boeing 737) is altitude H = 12 km, speed V = 240 m/s, specific fuel consumption Cs = 0.00019 kg fuel/s/kg thrust. Any deviation from this flight regime significantly reduces the maximum range (by up to 10–50%). The vehicle with a variable wing area loses 50% less range than a vehicle with a fixed wing

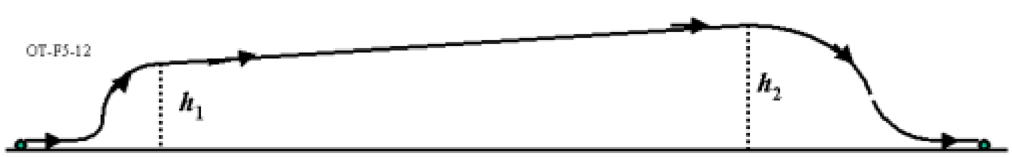


Fig. A4.22. Optimal trajectory of aircraft.

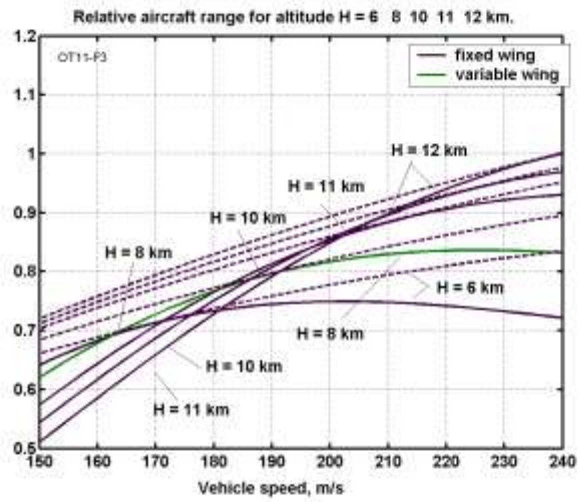


Fig. A4.23. Relative aircraft range for altitude H = 6, 8, 10, 11 and 12 km, maximum range R_m = 4381 km, relative fuel mass M_r = 0.2, body drag coefficient C_b = 0.08, wing drag coefficient C_d = 0.02.

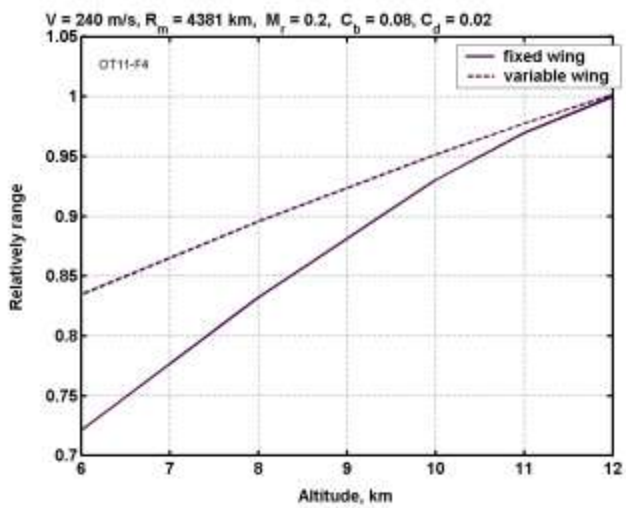


Fig. A4.23. Relative aircraft range for speed V = 240 m/s, maximum range R_m = 4381 km, relative fuel mass M_r = 0.2, body drag coefficient C_b = 0.08, wing drag coefficient C_d = 0.02.

General discussion and conclusion

a) The current space missiles were designed 30–40 years ago. In the past we did not have navigation satellites that allowed one to locate a missile (warhead) as close as 1 m to a target. Missile designers used inertial navigation systems for ballistic trajectories only. At the present time, we have a satellite navigation system and cheap devices, that enable aircraft, sea ships, cars, vehicles, and people to be located. If we exchange the conventional warhead for a warhead with a simple fixed wing with having a control and navigation system, we can increase the range of our old rockets 4.5–5 times (Fig. A4.13) or significantly increase the useful warhead weight (Fig. A4.19). We can also notably improve the precision of our aiming.

b) Current artillery projectiles for big guns and cannons were created many years ago. The designers assumed that the observer could see an aim point and correct the artillery. Now we have a satellite navigation system that allows one to determine the exact coordinates of targets and we have cheap and light navigation and control devices that can be placed in the cannon projectiles. If we replace our cannon ballistic projectiles with projectiles with a fixed wing, and a control and navigation system, we increase the range 3–9 times (from 35 km up to 360 km, see Fig. A4.20, A4.21). We can use a smooth barrel to increase the nozzle shell speed up to 2000 m/s and range up to 1000 km. These systems can guide the **winged** projectiles and significantly improving their aim. We can reach this result because we use all the **kinetic** energy of the projectile. A conventional projectile cannot remain in the atmosphere and drops at a very high speed. Most of its kinetic energy is wasted. In our case 70–85% of the projectile's kinetic energy is used for support of the moving projectile. This way the projectile range increases 3–9 times or more.

c) All aircraft are designed for only one optimal flight regime (speed, altitude, and fuel consumption). Any deviation from this regime decreases the aircraft range. For aircraft like to the Boeing 747 this regime is: altitude H = 12 km, speed V = 240 m/s, specific fuel consumption Cs = 0.00019 kgf/s/kg thrust. If the speed is reduced from 240 m/s to 200 m/s, the range decreases by15% (Fig. A4.23). Application of the variable wing area reduces this loss from 15% to 10%. If the aircraft reduces its altitude from 12 km to 9 km, it loses 12% of its maximum range (Fig. A4.24). If it has a variable wing area, it loses only 7.5% of its maximum range. Civil air vehicles are forced to deviate from the optimal conditions by weather or a given flight air corridor. Military air vehicles sometimes have to make a very large deviation from the optimal conditions (for example, when they fly at low altitude, below the enemy radar system). A variable wing area may be very useful for them because it decreases the loss by approximately 50%, improves supersonic flight and taking off and landing lengths.

The author offers some fixed and variable wing designs for air vehicles (Fig. A4.25). Variants **a, b, c**, and **f** are for missiles and warheads, variants **d,** and **e** are for shells

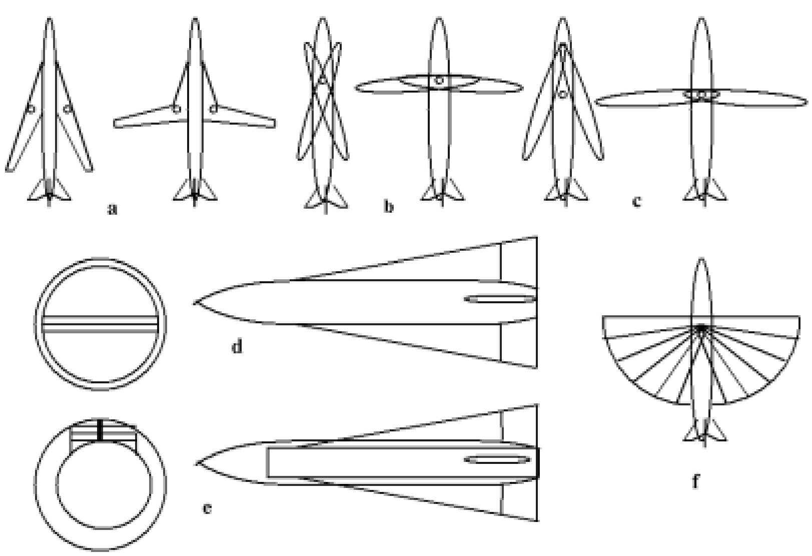


Fig. A4.25. Possible variants of variable wing designs: a, b, c, and f for aircraft; d and e for gun projectiles.

References for Attachment 1

1. A. Miele, "General Variable Theory of the Flight Paths of Rocket-Powered Aircraft, Missile and Satellite Carries", *Astronautica acta*, Vol. 4, pp. 256–288, 1958.
2. A.A. Bolonkin, "Special Extrema in Optimal Control Problems", Akademiya Nauk, Izvestia, *Theknicheskaya Kibernetika*, No. 2, March–April, 1969, pp. 187–198. See also English translation in *Eng. Cybernetics*, No. 2, March–April 1969, pp. 170–183.
3. A.A. Bolonkin*, Book:"***New Methods of Optimization and their Applications"**. Moscow, **MSTU. (Russian). «Новые методы Оптимизации и их применение» , Москва, МВТУ имени Баумана, 220 стр. (См. РГБ, Российская Государственная Библиотека, Ф-801-83/869-6). http://vixra.org/abs/1504.0011 v4.**
4. A.A. Bolonkin, "A New Approach to Finding a Global Optimum", *New Americans Collected Scientific Report*s, Vol. 1, 1991, The Bnai Zion Scientists Division, New York.
5. A.A. Bolonkin, "Optimization of Trajectory of Multistage Flight Vehicles", in collection *Researches of Flight Dynamics*, Masinobilding, Moscow, 1965, pp. 20–78 (in Russian).
6. A.A. Bolonkin, N.S. Khot, "Method for Finding a Global Minimum", AIAA/NASA/USAF/SSMO Symposium on Multidisciplinary Analysis and Optimization, Panama City, Florida, USA, 7–9 Sept. 1994.
7. A.A. Bolonkin, "Solution Methods for Boundary-Value Problems of Optimal Control Theory". 1973. Consultants Bureau, a division of Plenum Publishing Corporation, NY. Translation from *Prikladnaya Mekhanika*, Vol. 7, No 6, 1971, p. 639–650 (English).
8. A.A. Bolonkin, "Solution General Linear Optimal Problem with One Control". Journal "*Pricladnaya Mechanica*", Vol. 4, No 4, 1968, pps. 111–122, Moscow (in Russian).
9. A.A. Bolonkin, N. Khot, "Optimal Structural Control Design", IAF-94-1.4.206, 45th Congress of the International Astronautical Federation, World Space Congress-1994, 9–14 October, 1994, Jerusalem, Israel.
10. A. A. Bolonkin, N. S. Khot, "Optimal Bounded Control Design for Vibration Suppression", *Acta Astronautica*, Vol. 38, No. 10, pp. 803–813, 1996.
11. A.A. Bolonkin, D. Jeffcoat, "Optimal Thrust Angle of Aircraft", 2002, see Attachment 6.
12. A.A. Bolonkin, R. Sierakowski, "Design of Optimal Regulators", 2nd AIAA "Unmanned Unlimited" Systems, Technologies, and Operations – Aerospace, Land, and sea Conference and Workshop & Exhibit, San Diego, California, 15-18 Sep 2003, AIAA-2003-6638.
13. A.A. Bolonkin, Optimal Trajectories of Air and Space Vehicles", *Aircraft Engineering and Aerospace Technology*, **Vol.78, No.2,** pp.192–214.
14. Bolonkin's site: **http://Bolonkin.narod.ru** .
15. A.A. Bolonkin, Post-Doctoral dissertation "New Methods of Optimization and their Applicatuon". Moscow, Sant-Petersbursky Technical Universitete, 285 ps, (in Russian). Новые методы оптимизации и их применение. Диссертация на соискание ученой степени доктора технических наук. Москва. ЛПИ, 285 стр.
16. A.A. Bolonkin, Synopsis of post PhD dissertation: "New Methods of Optimization and their Applicatuon". Moscow, (Russian). Новые методы оптимизации и их примеение в задачах динамики управляемых систем. (Russian) . Автореферат диссертации на соискание ученой степени доктора технических наук. Москва, ЛПИ, 28 стр. http://viXra.org/abs/1503.0081

==

Attachment 2

Impulse solutions in optimization problems

Alexander Bolonkin

C&R, abolonkin@gmail.com

Abstract

The author considers the optimization problem named 'the impulse regime', when the control can have for a short time an instantaneous infinity value and the phase variables have gaps. In mathematics these mean: the variables are not continuous, not differentiable. The variable calculation and Pontryagin principle are not applicable. These problems are in space trajectories, theory of corrections, nuclear physics, economics, advertising and other real control tasks. We need a special theory and special methods for solution of these problems.

Author offers the following method, which simplifies and solves these tasks.

Introduction

Optimization methos are widely used in solving of technical problems. However, there are important classes of problems where they have great difficulties in the application. For example, in problems of space travel. The fact that the operational time of conventional rocket liquid propulsion is small (minutes), while the passive time of the interplanetary flight is large (months). In the result, we can consider the rocket work as an impulse, the speed as a jump which must expend minimum fuel. In mathematics, this means: the control is at an infinity value, the phase variables have a gap, and the variables are not continuous, not differentiable. The variable culculation and the Pontryagin principle are not applicable.

In 1968 the author offered the special methods [1] (see also [2 – 3]) for solution of the difference cases the impulse regime. In book [4] he applied this method to aerospace problems. Authors of work [5] developed the impulse theory for a particular case (linear version of control) using the theory of δ-functions. But his solutions are very complex and not acceptable in many practical problems.

In the given article the author offers a simpler method for solution of these problems: he shows the known impulse problems can be reduced to the special Pontryagin problem. Solution of them may be simple when methods are used in present time.

Statement of the problem

1. Statement of the conventional Optimization Problem. Assume the state of system is described by conventional differential equations:

$$I = F[x(t_1), x(t_2)] + \int_{t_1}^{t_2} f_0(t,x,u)dt, \quad \dot{x}_i = f_i(t,x,u) \quad i = 1,2,...,n, \quad (1)$$

where I is the objective function, x is n – dimentinal continuous piece-difference function of phase coordinates; u is r – dimentional piece-continuous, piece-diffrence functions of control, $a_i \le u_i \le b_i$, $i = 1,2,\ldots,r$, a,b = const; t is time. End values of $x(t_1)$, $x(t_2)$ are given or mobile. F is function of the end values $x(t)$.

We must find the control u, which gives the minumum the objective function I.

In our case (impulse problem) the control (or some its components) is at infinity (a very short time), the some (or all) phase variables have the gaps, and the variables are not continuous, not differentiable. The variable calculation and Pontryagin principle are not applicable.

2. Impulse Optimization Problem. Method of Solution.

The author offer the following method for solution of impulse problems.

We enter the special constants (unknown limited values) of impulses (sums, gaps)

$$v_i \quad i = 1,2,...,m. \quad (2)$$

These values may be binded the contions

$$x_i^+ = x_i^- + v_i, \quad \varphi_i(t,x,u,v) = 0 \quad i = 1,2,...,s, \quad s < m \quad (3)$$

and limitations

$$c_{i,1} \le v_i \le c_{im2} \quad i = 1,2,...,m, \quad (4)$$

where x_i^-, x_i^+ are x_i is phase coordinate on left and on right from point of impulse (gap), $c_{i,1}, c_{i,2}$ are consts. In particule, v can be unknown constant or zero.

The optimal problem is written in fo

$$I = F[x(t_1), x(t_2)] + \int_{t_1}^{t_2} f_0(t,x,u,v)dt, \quad \dot{x}_i = f_i(t,x,u,v) \quad i = 1,2,...,n,$$
$$\varphi_i(t,x,u,v) = 0, \quad i = 1,2,...,m \quad (5)$$

Where v are unknown limited impulses (gaps). End values of $x(t_1)$, $x(t_2)$ are given or mobile.

According [2], [3], we can write the generalized functionality introdused in form

$$J = I + \alpha, \quad (6)$$

where *J* - the generalized functionality introduced in [2],[3] p. 42, α is so named α – function introduced in [2],[3] (function equals zero on acceptable set, for example, on curves satisfying the equations (1) – (4)).

In our case we take

$$\alpha = \int_{t_1}^{t_2} \left[\sum_{i=1}^{i=n} \lambda_i(t,x)[\dot{x} - f_i(t,x,u,v)] + \sum_{i=n+i}^{i=n+m} \lambda_i(t,x)\varphi_i(t,x,u,v) \right] dt \quad (7)$$

Where $\lambda(t,x)$ is an unknown vector function.

We can re-write (6) as (see [3] p.42)

$$J = I + \alpha = A + \int_{t_1}^{t_2} B dt, \quad (8)$$

where (for brevity repeated indices are summed):

$$A = F + \lambda_i x_i \big|_{t_1}^{t_2}, \quad B = f_0 - \left(x_j \frac{\partial \lambda_j}{\partial x_i} + \lambda_i \right) f_i - x_i \frac{\partial \lambda_i}{\partial t}, \quad (9)$$

From Theorem 3.8 [3] we get: if we find at least one solution of particular equation about λ

$$J = \inf A + \inf B, \quad \inf_{u,v} \left[f_0 - \left(x_j \frac{\partial \lambda_j}{\partial x_i} + \lambda_i \right) f_i - x_i \frac{\partial \lambda_i}{\partial t} \right], \quad \frac{\partial B}{\partial x} = 0, \quad (10)$$

for the end condition inf *A*, we get optimal solution.
Note, the *B* (9) is different from the well-known Gamiltonian. If we will take the different function $\lambda(t,x)$, we will get the different conjugated system of equations $\partial B/\partial x = 0$.

In particular, if we will get $\lambda(t)$ ONLY as function *t*, we get the conventional Pontryagin principle of maximum

$$J = A + \int_{t_1}^{t_2} B dt, \quad (11)$$

where

$$A = F + \sum_{i=1}^{i=n} [\lambda_i(t_2) x_i(t_2) - \lambda_i(t_1) x_i(t_1)], \quad (12)$$

$$B = f_0(t,x,u,v) - \sum_{i=1}^{i=n} \lambda_i(t) f_i(t,x,u,v) - \sum_{i=n+i}^{i=n+m} \lambda_i(t) \varphi_i(t,x,u,v) \quad (13)$$

and equations

$$\dot{x}_i = f_i(t,x,u,v)\,,\quad i=1,2,...,n \tag{14}$$

$$\dot{\lambda}_i = \frac{\partial B}{\partial x_i}\quad i=1,2,...,n\,,\quad \inf_u B \quad \text{or}\quad \frac{\partial B}{\partial u_i}=0\quad i=,2,...,r,$$

$$\inf_v B \quad \text{or}\quad \frac{\partial B}{\partial v_i}=0\quad i=r+1,2,...,m, \tag{15}$$

The equations

$$\frac{\partial B}{\partial u}=0,\quad \frac{\partial B}{\partial v}=0 \tag{16}$$

are used only in the open area. λ_i are unknown multipliers.

Equations (11) - (16) gives the optimal trajectoris (minimum of I) of the system (5). We also must solve the boundary value problem – find such $\lambda_i(t_1)$ that to get the given $x_i(t_2)$.

The gap time t_θ and gap v inside interval ($t_1 < t_\theta < t_2$) we can also find the next way. Write the objective function in form

$$I = F[x(t_1),x(t_2)] + \Phi(t_\theta,x_\theta) + \int_{t_1}^{t_2} f_0(t,x,u)dt,$$
$$\dot{x}_i = f_i(t,x,u)\quad i=1,2,...,n, \tag{17}$$

where Φ is additional condition in t_θ (if they are given).
Write the general function as the sum of two functions in (t_1, t_θ) and ($t_1<t_\theta<t_2$)

$$J = F + \psi_2 - \psi_1 + \Phi + \psi_\theta^+ - \psi_\theta^- + \int_{t_1}^{t_\theta} Bdt + \int_{t_\theta}^{t_2} Bdt\,,$$
$$\text{where}\quad \psi_\theta(t_\theta) = \lambda_i x_i\,,\quad \psi_2(t_2)=\lambda_i x_i\quad \psi_1(t_1)=\lambda_i x_i. \tag{18}$$

In t_θ the minimal condition are

$$\inf_{t_\theta,x_\theta}[\Phi(t_\theta,x_\theta)+\psi_\theta^+(t_\theta,x_\theta)-\psi_\theta^-(t_\theta,x_\theta)],\quad \inf_{x,u} B = 0\,. \tag{19}$$

Here up "-" and "+" are values from left and right from point t_θ.

Notes:

1. We can find in form (3) ONLY the phase coordinates which we can aproximate as the impulse (in short time we can change a large value – for example, the speed in long flight, agle of trajectory, laser excitation of atom and so on). We cannot pulse space, distance, time.

2. The λ_i of corresponding coorditate has a gap/jump in moment of impulse. The moment (time) of gap or new λ_i (at right side) we can find (in open area) from the second equation (16). We must also to check up the ends of the intersal $[t_1, t_2]$.

3. In some cases, the optimal value of gap we can find by the selection of v.

4. The λ_i of f_i are functions of t, the λ_i of φ_i are constants.

Example

Let us to consider the typical problem of space travel - transfer from one space orbit to other. Assume the space ship has circular Earth orbit having the radius r_1 and speed V_0 . We want to reach the ecliptic orbit having the maximal radius $r_2 > r_1$ and spend the minimum of fuel. The liquid rocket engine works some seconds, the space flight is some months. That way we can consider the rocket flight as pulse mode which instant change speed (gap the speed). Our task is to find minimal gap of speed (minimal inpulse) $v = \Delta V$, because the minimal gap of speed is equivalent of the minimal expenditure of the rocket fuel.

Our objective function

$$I = \int_0^t \Delta V dt \qquad (20)$$

The variables (speed V and radius r) of free space flight in the Earth gravitation field is binded by the Law of energy conservation (kinetic + potencial energy equils constant c):

$$\frac{mV^2}{2} - m\mu\left(\frac{1}{r_0} - \frac{1}{r}\right) = c \,, \quad \text{or} \quad V^2 = \mu\left(\frac{2}{r} - \frac{2}{r_1 + r_2}\right) , \qquad (21)$$

Where m is mass space ship (satellite) mass, kg; r_0 is initial radius, m; μ is gravity constant. For Earth $\mu = 3.9802 \cdot 10^{14}$ m^3/s^2 , for Sun $\mu = 1{,}3276 \cdot 10^{20}$ m^3/s^2. That is elliptic orbite, r_1 is the radius of perigee; r_2 is the radius of apogee. We want to arrive from the circular orbite having V_0 , the radius $r_0 = r_1$ (the point of perigee) to point of apology r_2.

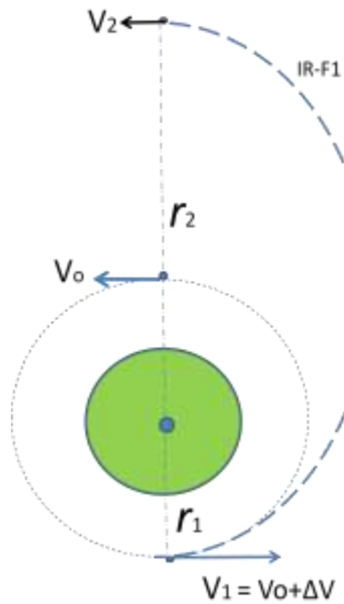


Fig.1. Orbite transver.

For elliptic orbits, the equation (21) may be re-writen in form:

$$(V_0+\Delta V)^2-2\mu\left(\frac{1}{r_1}-\frac{1}{r_1+r_2}\right)=0 \quad \text{or} \quad V_0+\Delta V-\sqrt{2\mu\left(\frac{1}{r_1}-\frac{1}{r_1+r_2}\right)}=0\ , \tag{22}$$

where; V_0 is speed on circular orbite having the radius r_1. The speed of circular orbit is

$$V_0=\sqrt{\frac{\mu}{r_0}}\ ,\quad V_1r_1=V_2r_2\,. \tag{23}$$

Here V_2 is speed in r_2. Last equation in (23) is Law of momentum conservation free flight in the central gravitation field.

Let us the write the function B (13) for left end in right side of point t_1 .

$$B=\Delta V+\lambda\left[V_1+\Delta V-\sqrt{2\mu\left(\frac{1}{r_1}-\frac{1}{r_1+r_2}\right)}\right], \tag{24}$$

From equation (16) we have

$$\frac{\partial B}{\partial(\Delta V)}=1+\lambda=0\ . \tag{25}$$

The equaetion (25) together with the equations (22),(23) allow to find the λ and the speed gap ΔV :

$$\lambda = -1, \quad \Delta V = \sqrt{\frac{\mu}{r_1}}\left(\sqrt{\frac{2r_2}{r_1+r_2}}-1\right) = V_0\left(\sqrt{\frac{2r_2}{r_1+r_2}}-1\right) = V_0\left(\sqrt{\frac{2\bar{r}}{\bar{r}+1}}-1\right) = V_0\left(\frac{V_1}{V_a}-1\right), \quad (26)$$

where

$$\bar{r} = \frac{r_2}{r_1}, \quad V_a = \sqrt{\frac{V_1^2+V_2^2}{2}}, \quad V_1 r_1 = V_2 r_2 . \quad (27)$$

Here V_2 is speed in apogee, V_a is average speed.
We reached the request r_2 by the first impulse. That way we don't need the additional impulse and reseach.

The formula (26) for computation ΔV is known as transfer in Gohman ellipse [6]. New is proof of optimization.

The reader can solve same way the more complex inpulse (gap) problems [4].

Referances

1. Bolonkin A.A., Impulse solutions in problem of the optimal control. Prosiding of Siberian Branch of the Academy of Sciences of the USSR, series of Technical sciences. #13, Issue 3, 1968 (in Russian).

2. Bolonkin A.A., New methods of Optimizations and their application. The post-doctoral Ph.D thesis, 1969, Chapter 5, pp.120-141. Leningrad Politechnic University (in Russian).

   http://viXra.org/abs/1509.0267 Part 1, http://vixra.org/abs/1509.0265 Part2, https://www.academia.edu/s/2a5a6f9321?source=link

3. Bolonkin A.A., New methods of Optimization and their applications. Moscow High Technical University, named Bauman, 1972. 220 ps (In Russian: Новые методы оптимизации и их применение. МВТУ им. Баумана, 1972г., 220 стр., См. РГБ, Российская Государственная Библиотека, Ф-861-83/1809-6**).** http://vixra.org/abs/1504.0011 v4**. ,** https://www.academia.edu/11054777/ v.4
4. Bolonkin AA., "Non Rocket Space Launch and Flight". Elsevier, 2005. 488 pgs. ISBN-13: 978-0-08044-731-5, ISBN-10: 0-080-44731-7 . http://vixra.org/abs/1407.0174 , https://archive.org/details/Non-rocketSpaceLaunchAndFlightv.3 , (v.3),
5. Дыхта В.А., Самсонюк О.Н., Оптимальные импульсные управления с приложениями. М.Наука,ФМ. 2000, (in Russian).
6. Walter Hohmann. Die Erreichbarkeit der Himmelskörper. — Verlag Oldenbourg in München 1925. — ISBN 3-486-23106-5.
7. Bolonkin A.A.,**Impulse solutions in optimization problems** 11 20 15. http://viXra.org/abs/1511.0189; https://www.academia.edu/s/3244c0c4f0?source=link ???

https://www.academia.edu/s/00538971c8 ,
https://archive.org/details/ArticleImpulseSolutionsdoc200311115AfterJoseph
GSJornal: http://gsjournal.net/Science-Journals/Research%20Papers-Astrophysics/Download/6259

30 September 2015.

==

Book
"Universal Optimization and its Application"

Chapter 12. Optimal Thrust Angle of Aircraft

Summary. The optimal angle for an aircraft's thrust vector is derived from first principles. Two equations are shown to encompass six different flight regimes. The main result for take-off and landing is that the optimal thrust angle in radians approximately equals the coefficient of rolling friction. For climb, cruise, turn and descent, the optimal thrust angle equals the arctangent of the ratio of drag coefficient to lift coefficient. The second result differs from the well-known result that optimal thrust angle equals the arctangent of the partial derivative of drag with respect to lift. The author discusses this difference

Nomenclature

B = artificial function, $\lambda dx/dt - H$

C_D = drag coefficient

C_L = lift coefficient

D = drag

d = take-off or landing distance

E = aircraft efficiency, C_L/C_D

F = fuel consumption

f = performance function

g = n-dimensional vector constraint function

g_0 = acceleration due to gravity

H = Hamiltonian, $-f + \lambda g$

h = altitude

I = performance index

L = lift force

M = aircraft mass

OTA = Optimal Thrust Angle

q = dynamic pressure, $\rho V^2 / 2$

R = range

S = wing area

t = time

T = thrust or time of flight

T_f = friction force

u = *m*-dimensional control vector

V = aircraft speed

W = aircraft weight

w_f = specific fuel consumption

x = *n*-dimensional state vector

λ = n-dimensional Lagrange multiplier

γ = thrust angle

μ = friction coefficient

ψ = specific function

ϕ = roll angle

Introduction

Aircraft designers must determine the angle of the thrust vector relative to the main horizontal flight direction. When this angle is positive (up from the horizontal plane), an additional lift force is generated, but at the expense of horizontal thrust. In this paper, the optimal thrust angle is derived, using both classical methods and an alternative optimization method developed by the first author.[1,2] Many methods of deflecting the nozzle exhaust stream of rocket engines to provide thrust vector control have been investigated, including jet vanes, gimbaled or swiveled nozzles, and extendable nozzle deflectors.[3,4,5,6] Jet vanes have been widely applied for the control of solid rocket engines and for early liquid-rocket engines, including the German V-2 missile.[7] Reference 8 presents metrics for assessing the performance of fighter aircraft implementing thrust vector control.

References 3 and 9 are most closely related to this paper. In Reference 3, the authors use numerical calculations to search for the optimal thrust angle, whereas in this paper the focus is theoretical, rather than numerical. In Reference 9, Miele presents a basic theory for analyzing the optimum flight paths of rocket-powered vehicles. Miele simultaneously optimizes the time history of lift, thrust modulus and thrust direction, and states that the optimal thrust angle equals the arctangent of the partial derivative of drag with respect to lift. In this paper, we provide theory and formulas for the OTA for six primary flight regimes of any aircraft type. The formulas provided are accurate for stable flight conditions, but may be sub-optimal during high dynamic maneuvers. The six flight regimes are listed below, each with one or more optimization objectives.

1. Take-off, to minimize take-off ground run.
2. Climb, to minimize fuel consumption.
3. Cruise, to minimize fuel consumption or to maximize range.
4. Turn, to minimize fuel consumption or to minimize turn time.
5. Descent, to minimize fuel consumption or to maximize range.
6. Landing, to minimize landing roll.

General Methodology

Consider the problem of minimizing a performance index *I*, where

$$I = \int_0^T f(t,x,u)dt \quad (1)$$

We wish to minimize *I* with respect to *x* and *u*, subject to the dynamic constraint

$$\dot{x} = g(t,x,u) \quad (2)$$

We assume an initial condition, $x(0)$, is known. Following the approach described in Chapters 1, §4B, 2 or in Reference 1, we define an artificial *α* - function in form

$$B = f - \frac{\partial \psi}{\partial x} g(t,x,u) - \frac{\partial \psi}{\partial t} \quad (3)$$

In particular, the function ψ may be defined by

$$\psi = \lambda(t) \cdot x \qquad (4)$$

If we find

$$\min_{x,u} B = \min_{x,u}[f(t,x,u) - \lambda(t)g(t,x,u) - \dot{\lambda}(t)x] \qquad (5)$$

then the values of *x* and *u* that minimize *B*, subject to the constraint given in Eq. (2), are optimal control and state vectors for the problem stated in Eq. (1).

For readers are not friendly with Method of Deformation, we can also solve this problem by a conventional method[10] using the Hamiltonian for this problem (as it is shown in Chapter 2 most conventional methods may be received from Method of Deformation), which is given by

$$H(t,x,u,\lambda) = -f(t,x,u) + \lambda(t) \cdot g(t,x,u), \qquad (6)$$

and

$$\dot{\lambda} = -\frac{\partial H}{\partial x}$$

As it is shown in Chapters 1 – 2, Hamiltonian is part of particular function *B*, which is particular case of *α* – function and *α* – function is particular case of *β* – function

$$\min_{x,u} B = \min_{x,u}[f(t,x,u) - \lambda(t)g(t,x,u) - \dot{\lambda}(t)x] = \min_{x,u}[-H - \dot{\lambda}(t)x] \ .$$

In *B* we find the minimum for *x, u*. In conventional method we find the maximum of *H* only for u.

The values of *u* which maximize the Hamiltonian, subject to the constraint in (2), are optimal control vectors for (1). That is

$$\overline{H} = \max_{u} H(t,x,u,\lambda) \ . \qquad (7)$$

When the process does not change with time, we have a more straightforward problem:

Minimize a performance index *I*, defined by

$$I = \int_0^T f(x,u)dt \qquad (8)$$

with respect to *x* and *u*, subject to the dynamic constraint

$$\dot{x} = g_i(x,u), \quad \text{for } i = 1,2, \ldots, n \tag{9}$$

$$H(t,x,u,\lambda) = -f(x,u) + \lambda(t)g(x,u) \tag{10}$$

$$\overline{H} = \max_u H(t,x,u,\lambda), \qquad \overline{u} = u(t,x) \tag{11}$$

The parameter λ is an *n*-dimensional unknown Lagrange multiplier and $\overline{u}$ is the optimal control. Eqs. (4) through (6) give the system of equations

$$\frac{\partial B}{\partial u_j} = \frac{\partial H}{\partial u_j} = 0\,, \quad j = 1,2, \ldots, m; \quad \frac{\partial B}{\partial x_i} = \dot{\lambda}_i(t) + \frac{\partial H}{\partial x_i} = 0\,, \quad i = 1,2, \ldots, n \tag{12}$$

These equations are equivalent to conventional principle of maximum[10];

$$\dot{\lambda}_i(t) = -\frac{\partial H}{\partial x_i}\,, \quad i = 1,2,\ldots, n; \qquad \frac{\partial H}{\partial u_j} = 0\,, \quad j = 1,2,\ldots,m. \tag{13}$$

These equations, together with Eq. (2), allow us to find an extreme of the Hamiltonian *H*, which is optimal if the appropriate second order sufficient conditions for optimality are satisfied.

Optimal Thrust Angle for Take-off and Landing

For take-off, the performance index is the take-off distance, described by

$$d = \int_0^T V dt\ . \tag{14}$$

The aircraft speed serves as the performance function. The dynamic constraint on acceleration is given by

$$\dot{V} = \frac{1}{M}(T\cos\gamma - D - T_f) \tag{15}$$

as illustrated in Fig. 1.

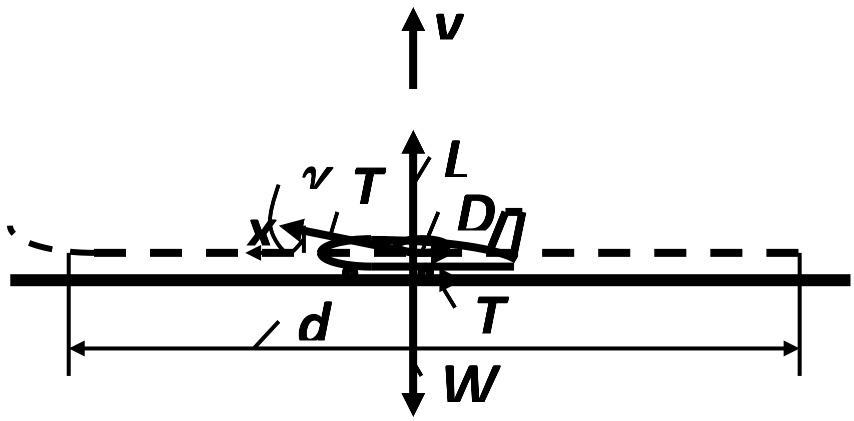


Fig. 1 Take-off

The friction force is given by

$$T_f = \mu \cdot (W \cdot g_0 - L - T \sin\gamma) \tag{16}$$

We know from aerodynamics and trigonometry that

$$L = C_L qS, \qquad D = C_D qS, \qquad \sin\gamma = \sqrt{1-\cos^2\gamma} \tag{17}$$

We consider only the positive root, but the result is the same for the negative root.

To simplify subsequent calculations, make the substitution

$$u = \cos\gamma \tag{18}$$

so that

$$\sin\gamma = \sqrt{1-u^2} \tag{19}$$

Substituting Eqs. (16) through (19) into (15) yields

$$\dot{V} = \frac{1}{M}[T \cdot u - D - \mu \cdot (W \cdot g_0 - L - T \cdot \sqrt{1-u^2})] \tag{20}$$

which leads to function *B* or the Hamiltonian *H*

$$B = f + \dot{\lambda}(t) \cdot V + \lambda(t) \cdot \dot{V} = V + \dot{\lambda}(t) \cdot V + \lambda(t) \cdot \frac{1}{M}[T \cdot u - D - \mu \cdot (W \cdot g_0 - L - T\sqrt{1-u^2})]$$

$$= \dot{\lambda}(t) \cdot V + H \ , \tag{21}$$

where $$H = V + \lambda(t)\frac{1}{M}[T \cdot u - D - \mu \cdot (W \cdot g_0 - L - T\sqrt{1-u^2})] \tag{22}$$

To find the minimum of *B* over all admissible *u*, the necessary condition is that the partial derivative is equal to zero, that is,

$$\frac{\partial B}{\partial u} = 0 \tag{23}$$

or

$$\frac{\partial B}{\partial u} = \frac{\lambda(t) \cdot T}{M}\left[1 - \frac{\mu \cdot u}{\sqrt{1-u^2}}\right] = 0 \tag{24}$$

If *M, T*, and $\lambda \neq 0$, then from (24), it must be true that

$$\mu \cdot u = \sqrt{1-u^2} \tag{25}$$

or

$$\mu^2 \cdot u^2 = 1 - u^2 \tag{26}$$

so that the final result from (26) is

$$u = \pm\frac{1}{\sqrt{1+\mu^2}} \tag{27}$$

Returning to the original notation, we have the thrust angle as a function of the coefficient of friction,

$$\cos\gamma = \pm\frac{1}{\sqrt{1+\mu^2}} \quad \text{or} \quad \gamma = \cos^{-1}\left(\frac{\pm 1}{\sqrt{1+\mu^2}}\right) \text{ radians .} \tag{28}$$

We can use the trigonometric identity

$$\cos\gamma = \frac{1}{\sqrt{1+\tan^2\gamma}} \tag{29}$$

to get our final result,

$$\boxed{\tan\gamma = \pm\mu} \tag{30}$$

or, for small μ, say μ < 0.2, we have the design rule-of-thumb that

$$\boxed{\gamma \cong \pm\mu} \tag{31}$$

where γ is in radians.

The sign of γ depends on our goal, minimization or maximization of the function, as well as the sign of λ and *T* in Eq. (24). Clearly, the thrust must have a forward direction for aircraft take-off, and the angle γ must be positive. Similarly, for landing, the thrust must have a backward direction to brake the airplane, and the angle γ must be negative, pushing the airplane to the ground, as illustrated in Fig. 2.

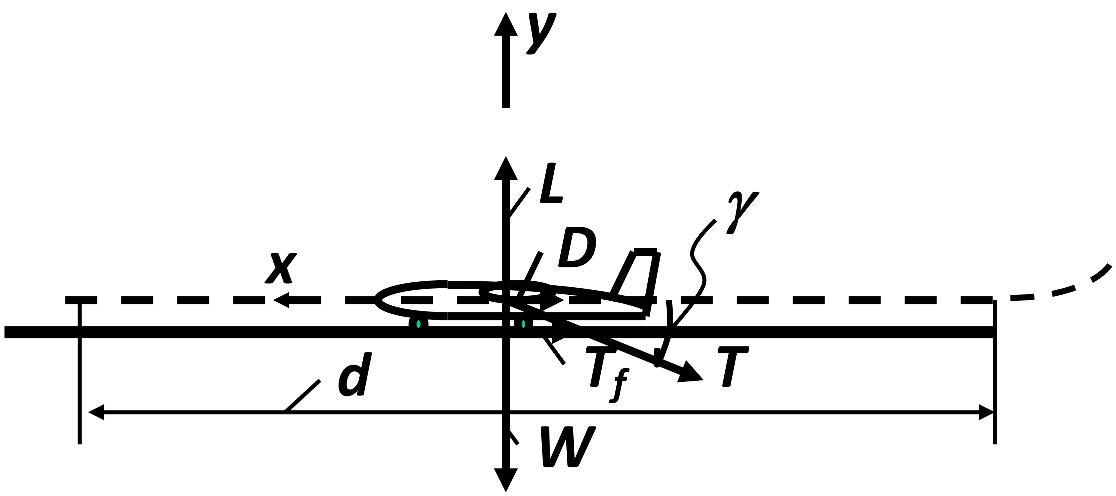


Fig. 2 Landing

The angles for take-off and landing are different, because the coefficients of rolling friction are different for take-off and landing. For take-off, the friction coefficient is small (μ approximately 0.01 – 0.05); for landing, the coefficient is larger (μ approximately 0.3 – 0.4). The direction of thrust is also different for take-off ($\gamma \cong$ +1 to +3 degrees) than for landing ($\gamma \cong$ -16 to -22 degrees). For take-off, the thrust has a forward direction; for landing the thrust has a backward direction. As a design "rule-of-thumb," we can say that the OTA in radians is equal to the coefficient of rolling friction for take-off, and the OTA is within 5% of the coefficient of rolling friction for landing. The expression tan γ = μ is exact for any rolling friction coefficient.

The optimal angles for take-off and landing are graphed in Figs. 3 and 4, respectively.

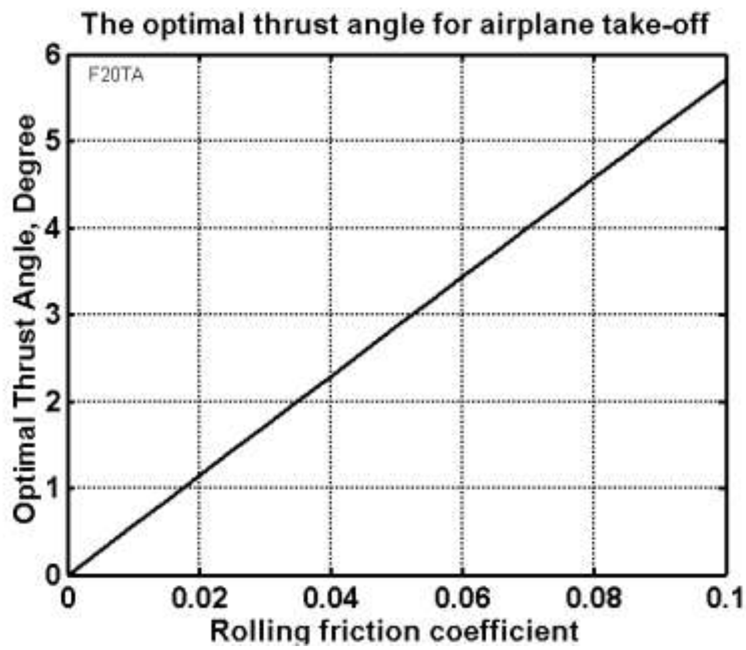


Fig. 3. Optimal Thrust Angle for Take-off

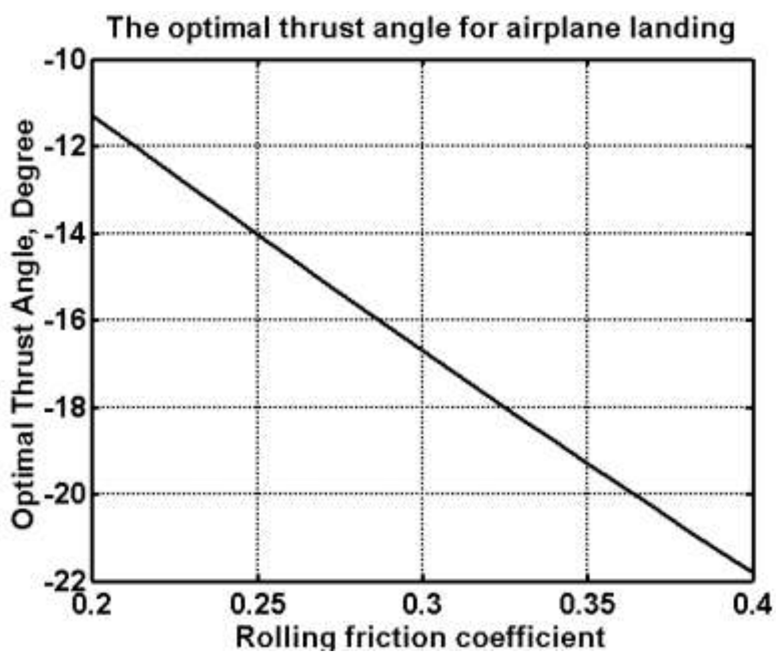


Fig. 4. Optimal Thrust Angle for Landing

Optimal Angle of Thrust Vector in Horizontal Flight (Cruise Regime)

Assume that speed, altitude, and direction of flight are constant during horizontal flight time, and that we wish to maximize range, *R*, of the aircraft over the time interval [*0,T*]. Then

$$R = \int_0^T V dt \qquad (32)$$

The equilibrium equations of motion (Fig. 5) are

$$T\cos\gamma - D = 0\,, \qquad (33)$$

$$L - W \cdot g_0 + T\sin\gamma = 0 \qquad (34)$$

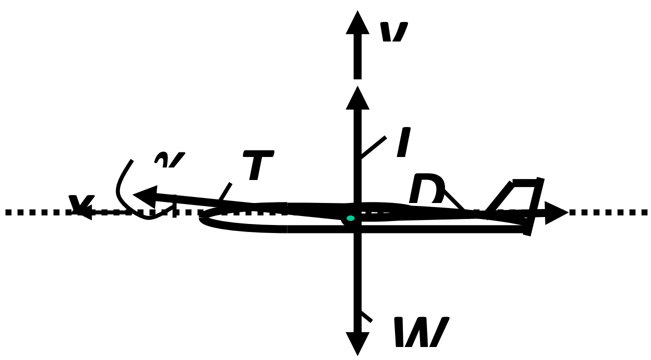


Fig. 5. Horizontal Flight.

Using the notation

$$E = \frac{L}{D} = \frac{C_L}{C_D}\,, \qquad u = \cos\gamma\,, \qquad \text{and } \sin\gamma = \sqrt{1-u^2} = 0 \qquad (35)$$

We can substitute *D* and *u* from (35) and *L* from (34) into (33) to obtain

$$T \cdot u - \frac{W \cdot g_0}{E} + \frac{T}{E}\sqrt{1-u^2} = 0 \qquad (36)$$

Next, compose the Hamiltonian function *H,* as in Eq. (10)

$$H = -V + \lambda \cdot (T \cdot u - \frac{W \cdot g_0}{E} + \frac{T}{E}\sqrt{1-u^2}) \qquad (37)$$

And find the maximum of this function

$$\frac{\partial H}{\partial u} = \lambda \cdot T\left[1 - \frac{u}{E\sqrt{1-u^2}}\right] = 0 \tag{38}$$

If we take values for λ and T such that $\lambda \cdot T \neq 0$, we find that

$$u = E\sqrt{1-u^2} \tag{39}$$

or

$$u^2 = E^2\left(1-u^2\right) \tag{40}$$

From (40), it follows that

$$u = \pm \frac{E}{\sqrt{1+E^2}} \tag{41}$$

or

$$\cos\gamma = \pm \frac{E}{\sqrt{1+E^2}} \qquad \text{or} \qquad \gamma = \arccos\left(\pm \frac{E}{\sqrt{1+E^2}}\right) \tag{42}$$

Note that γ in degrees given by

$$\gamma^\circ = 180 \cdot \gamma / \pi \, . \tag{43}$$

From physical conditions, it is evident that angle γ is positive. For fighter aircraft, aerodynamic efficiency, *E*, ranges from two to ten. For transport or passenger aircraft, efficiency ratios vary from ten to twenty. Using the trigonometric identify in (29), we get a final result in simpler form,

$$\tan\gamma = \frac{1}{E} \qquad \text{or} \qquad \boxed{\tan\gamma = \frac{C_D}{C_L}} \tag{44}$$

For small γ, say $\gamma < 0.2$ radians, we have the design rule-of-thumb that

$$\boxed{\gamma \cong \frac{C_D}{C_L}} \tag{45}$$

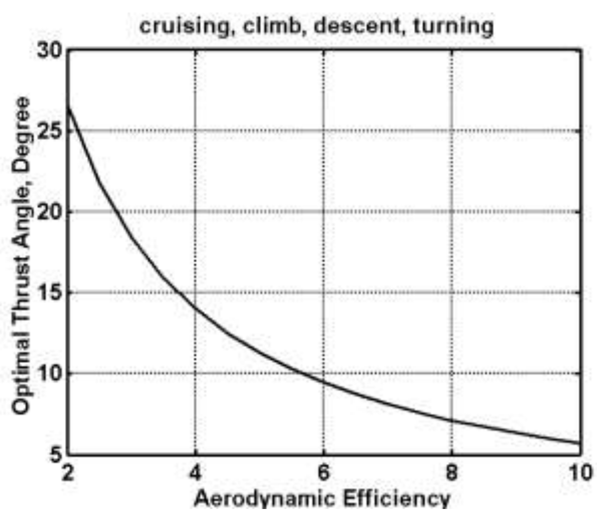
The OTA for the cruise regime is graphed in Fig. 6 for aerodynamic efficiencies ranging from two to ten, typical of fighter aircraft.

So far, we have used as the performance index the maximum range of the aircraft. The results are the same if we minimize fuel consumption,

$$F = \int_0^T w_f \cdot dt \qquad (46)$$

Or minimize time

$$T = \int_0^T dt \ . \qquad (47)$$



Thrust Angle for Cruise Regime

Climb and Descent Regimes

Fig. 7 illustrates the climb regime.

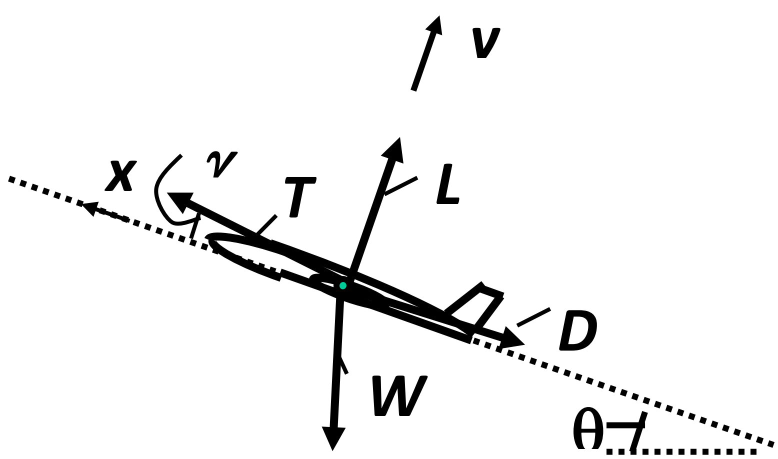


Fig. 7. Climb.

Let us take as the performance index the range or altitude,

$$R = \int_0^T V dt \qquad \text{or} \qquad h = \int_0^T V \sin\theta dt \tag{48}$$

Then the equilibrium equations are

$$T\cos\gamma - D - W \cdot g_0 \sin\theta = 0 \tag{49}$$

$$L + T\sin\gamma - W \cdot g_0 \cos\theta = 0 \tag{50}$$

where θ is the angle between trajectory and horizon. Using the notation

$$E = \frac{L}{D} = \frac{C_L}{C_D}, \qquad \cos\gamma = u\,, \qquad \sin\gamma = \sqrt{1-u^2}\,. \tag{51}$$

and substituting Eqs. (50) and (51) into Eq. (49), we have

$$Tu - \frac{W \cdot g_0 \cos\theta}{E} + \frac{T}{E}\sqrt{1-u^2} - W\sin\theta = 0 \tag{52}$$

And

$$H = V + \lambda[Tu - \frac{W \cdot g_0 \cos\theta}{E} + \frac{T}{E}\sqrt{1-u^2} - W \cdot g_0 \sin\theta] \tag{53}$$

The necessary condition for an extreme is

$$\frac{\partial H}{\partial u} = 0 \tag{54}$$

or

$$\frac{\partial H}{\partial u} = \lambda T\left(1 - \frac{u}{E\sqrt{1-u^2}}\right) = 0 \tag{55}$$

That is the same as Eq. (38), which means that the final equation for the optimal angle of thrust vector in climb and descent will be equal to the equation for a cruise regime.

$$\cos\gamma = \pm\frac{E}{\sqrt{1+E^2}} \qquad \text{or } \tan\gamma = \frac{1}{E} = \frac{C_D}{C_L} \qquad \text{or } \gamma \approx \frac{C_D}{C_L} \tag{56}$$

From physical conditions, it is evident that angle γ is positive. The aerodynamic efficiencies are different for climb, descent, and cruise, so that the optimal thrust vector angle will be different, but the equations for the calculation are the same. Again, note that we can use trigonometric equalities to derive the more concise expression, cot γ = E, which is exact for any aerodynamic efficiency ratio. The results are the same whether time or fuel consumption are used for the performance index.

Turning of airplane

Consider now the turning of an airplane in one plane, with a constant roll angle ϕ. Our performance index can be distance, minimum time of turn, or fuel consumption.

$$R = \int_0^T V dt, \qquad T = \int_0^T dt, \qquad F = \int_0^T w_f \, dt \tag{57}$$

The equations of motion are

$$T \cos - D = 0 \tag{58}$$

$$L + T \sin\gamma - W \cdot g_0 \cos\phi = 0 \tag{59}$$

Using the notation

$$E = \frac{L}{D} = \frac{C_L}{C_D}; \qquad \cos\gamma = u; \qquad \sin\gamma = \sqrt{1-u^2} \tag{60}$$

and substituting (59) and (60) into (58), we get

$$T \cdot u - \frac{W \cdot g_0 \cos\phi}{E} + \frac{T\sqrt{1-u^2}}{E} = 0 \tag{61}$$

and

$$H = V + \lambda \left[Tu - \frac{W \cdot g_0 \cos\phi}{E} + \frac{T}{E}\sqrt{1-u^2} \right] \tag{62}$$

The necessary condition for an extreme is

$$\frac{\partial H}{\partial u} = 0 \quad \text{or} \quad \frac{\partial H}{\partial u} = \lambda T \left(1 - \frac{u}{E\sqrt{1-u^2}} \right) = 0 \tag{63}$$

Eq. (63) is equivalent to Eq. (38), which means the final equation for the optimal angle of thrust vector in a roll is equal to the equation for a cruise regime

$$\cos\gamma = \pm \frac{E}{\sqrt{1+E^2}} \qquad \text{or } \tan\gamma = \frac{1}{E} = \frac{C_D}{C_L} \qquad \text{or } \gamma \approx \frac{C_D}{C_L} \tag{64}$$

From physical conditions, it is evident that angle γ is positive.

Discussion

The problem of determining an OTA is also discussed in Reference 9, in which the OTA for rocket-powered aircraft is given by

$$\omega = \arctan \frac{\partial D}{\partial L} \tag{65}$$

where ω is equivalent to our angle γ. In the particular case of a parabolic polar drag coefficient of the form $C_D = C_{D_0}(M) + K(M)C_L^2$, where M is the Mach number, K is the induced drag factor, and C_{D_0} is the zero-lift drag coefficient, Eq. (65) leads to

$$\omega = \arctan(2KC_L) \tag{66}$$

Eqs. (44) and (66) give very different results (Fig. 8). For example, when there is no lift force (C_L = 0), Eq. (44) gives γ = 90º, meaning that the optimal thrust angle is strictly vertical (perpendicular to the desired trajectory), while Eq. (66) gives ω = 0, corresponding to a horizontal thrust. Conversely, when the lift force is maximum, Eq. (65) gives ω = 90º. We also see in Fig. 8 that as the lift force (C_L) decreases after passing through its maximum point, Eq. (65) yields an optimal thrust angle greater than 90º, producing a reverse thrust force. So, we conclude that Eqs. (65) and (66) do not adequately model the OTA near extreme points.

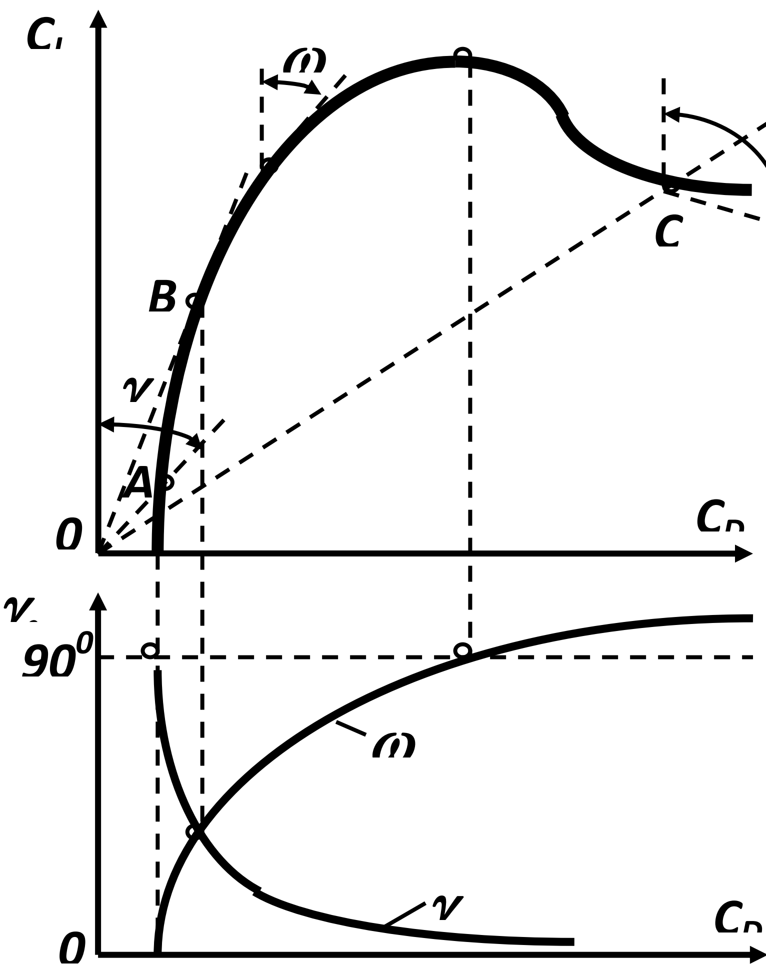


Fig. 8. Comparison of γ and ω .

The angle γ produced by Eq. (44) also has better trend characteristics, starting at 90º when $C_L = 0$, then decreasing as the lift force is increasing and positive. The angle ω starts at zero when $C_L = 0$, then increases as the aerodynamic lift force increases. Eqs. (44) and (65) do produce the same result at one point, when the efficiency coefficient, $E = C_L / C_D$ is maximized. Eq. (65) was derived for an optimal angle of attack, and the result is valid at the point of optimal aircraft lift. Eq. (44) is more general, and may be used at any polar coordinate in any of the four flight regimes: climb, cruise, turn or descent.

Conclusions

In this chapter, we derived two simple equations for the optimal thrust angle of an aircraft. One equation is valid for take-off and landing, the other for climb, cruise, turn, and descent. During take-off, the OTA is positive, decreases as the coefficient of rolling friction decreases, and is essentially equal to the friction coefficient. During landing, the OTA is negative, increases as the coefficient of rolling friction increases, and is within five percent of the value of the friction coefficient. The simple expression tan OTA = μ provides an exact result for the OTA as a function of the rolling friction coefficient.

In the climb, cruise, turn or descent flight regimes, the OTA depends only on the coefficient of aerodynamic efficiency. Here we observe an inverse proportion: the greater the coefficient of aerodynamic efficiency, the smaller the OTA. The OTA is positive in all flight regimes, with the possible exception of air braking, which is not addressed in this research. As in the cases of take-off and landing, we have a simple expression, tan OTA = 1 / E, relating the optimal thrust angle to a single parameter, the aerodynamic efficiency. The equations for OTA developed in this paper were also shown to have more intuitive trends and better behavior at extreme points than the Miele equations.

References to Chapter 13

[1] Bolonkin, Alexander, "A New Approach to Finding a Global Optimum," *New American's Collected Scientific Reports*, Vol. 1, The Bnai Zion Soviet-American Scientists Division, 1990.

[2] Bolonkin, Alexander, "New Methods of Optimization and their Application," Bauman Technical University, Moscow, 1972 (in Russian).

[3] Gilyard, Glenn and Bolonkin, Alexander, "Optimal Pitch Thrust-Vector Angle and Benefits for all Flight Regimes, NASA-TM-2000-209021, March 2000.

[4] Gal-Or, Benjamin, "Thrust Vectoring for Flight Control and Safety: A Review," *International Journal of Turbo and Jet Engines*, 11, 1994, pp. 119-136.

[5] Gerren, Donna S., "Design, Analysis, and Control of a Large Transport Aircraft Utilizing Selective Engine Thrust as a Backup System for the Primary Flight Control." NASA-CR-186035, September 1995.

[6] Mangold, P., and Wedekind, G, "Inflight Thrust Vectoring: A Further Degree of Freedom in the Aerodynamic/Flight Mechanical Design of Modern Fighter Aircraft," *AGARD, Aerodynamics of Combat Aircraft Controls and of Ground Effects*, April 1990.

[7] Koelle, Heinz Hermann, Ed., *Handbook of Astronautical Engineering*, McGraw-Hill, 1961, pp. 19-38.

[8] Kutschera, Antony, and Render Peter M., "Performance assessment of thrust vector controlled post stall manoeuvrable fighter aircraft using minimal input data," AIAA Paper 99-4020, 1999.

[9] Miele, Angelo, "General Variational Theory of the Flight Paths of Rocket-Powered Aircraft, Missiles and Satellite Carriers." *Astronautica acta*, Vol 4, New York: Pergamon Press, 1958, pp. 272-273.

[10] Pontryagin, L.S., V.G. Boltyanskii, R.V. Gamkrelidze, and E.F. Mischenko, *The Mathematical Theory of Optimal Processes*. New York: Interscience Publishers, Inc., 1962, pp. 17 - 20.

===

Book

"Universal Optimization and its Applications" by A.A. Bolonkin

Chapter 14. Design of Optimal Regulators

Summary

Current research suggests the use of a liner quadratic performance index for optimal control of regulators in various applications. Some examples include correcting the trajectory of rocket and air vehicles, vibration suppression of flexible structures, and airplane stability. In all these cases, the focus is in suppressing/decreasing system deviations rapidly. However, if one compares the Linear Quadratic Regulator (LQR) solution with optimal solutions (minimum time), it is seen that the LQR solution is less than optimal in some cases indeed (3-6) times that obtained using a minimum time solution. Moreover, the LQR solution is sometimes unacceptable in practice due to the fact that values of control extend beyond admissible limits and thus the designer must choose coefficients in the linear quadratic form, which are unknown.

The authors suggest methods which allow finding a quasi-optimal LQR solution with bounded control which is closed to the minimum time solution. They also remand the process of the minimum time decision.

Keywords: *Optimal regulator, minimum time controller, Linear Quadratic Regulator (LQR).*

The chapter is accepted as paper AIAA-2003-6638 by 2nd AIAA "Unmanned Unlimited" Systems, Technologies, and Operations- Aerospace, Land, and See Conference and Workshop & Exhibit, San Diego, California, USA, 15-18 Sep 2003.

Introduction

The LQR solution is easily and conveniently written using the Riccati equation as an optimal solution. The scientist who accepts this may be acting as an intoxicated man in a Russian anecdote: one night a man is observed creeping around a streetlight. A passerby asks him, what are you doing? – I lost money. Where did you lose the money? –There at the other end of the street. Then, why are you looking here? – This is where the light is!

The minimum time solution is more complex, however, it can be conveniently determined in many problems by the availability, generally, of high-speed computers. Also, this approaches us with a true minimum time solution.

For an *n*-dimensional problem with one control this solution found in general form in reference [1]. For the two-dimensional case this solution can be presented graphically, see ref. [1]. Methods for other general optimal solutions are offered in [2]-[4].

The LQR solution has three main issues:

1) The selection of the matrix coefficients in the performance index are designer selected and the solution is dependent upon the value of these coefficients.
2) The range of control values can be large in number and this not admissible for practice.
3) The “optimal” LQR solution can be up to 3-6 times worse, then the minimum time solution (see the example in this paper).

If a researcher chooses to use the LQR solution, the authors suggest a method for limiting maximum control (see point 2) as well as for the choice of selecting the coefficients in the performance index. This allows up to a 2-3 times improvement in the performance index (see accompanying examples) and thus makes the LQR solution acceptable in practical applications.

The traditional approach used in the design of a controlled structural system is to design the structure first by satisfying given requirements and then to design the control system. The structure is designed with such constraints placed on weight, allowable stresses, displacements, buckling, general instability, frequency distributions, etc. When the selection of the geometry, cross-sectional area of the members, and material are determined for a specified structure, then the structural frequencies and vibration modes become important input in the design of the control system. Some investigators have written papers discussing an integrated design approach for optimal control. In most references, the control design procedures used, do not take into consideration the limitations on the control forces developed by the actuators, and have not been treated as constraints or design variables. In this paper the problems associated with the selection of the performance index, parameters, weight coefficient in the LQR problem, and limitation of control forces are addressed.

In the following sections, theories for the synthesis of an optimal control laws with a quadratic performance index and bounded control forces are given. This is followed by a SISO (Single Input, Single Output) control problem designed using both approaches for comparison of the end state trajectories, with different bounds placed on control forces. Next, the control system for an idealized wing-box is used to illustrate a design application of the method. A discussion on the application of a control system with bounded control for an integrated design of a structure and control system can be found in ref. [5]. Related articles are [6]-[10].

1. Optimal Control

The general optimal control problem can be described by the following equations

$$I = F_0(x_1, x_2) + \int_{t_1}^{t_2} f_0(t, x, v)dt \; , \quad dx/dt = f(t,x,v), \quad x(t_1) = x_1, \quad x(t_2) = x_2 \qquad (1\text{-}1)$$

where I is the functional (objective function), t is time, x is a n-dimensional vector of state, and v is a p-dimensional vector of control forces. The vector $v \in V$ where V can be a bounded domain. Boundary conditions t_1, t_2, x_1, x_2 are usually given, $(t_1,t_2) \in T$.

The control parameter, v is calculated so that $I = min$. To find the solution to this problem by Method of Deformation of Chapter 2 (α – function), assume the function

$$\psi = \psi(t,x) \qquad (1\text{-}2)$$

and write the new function

$$J = A_1 + \int B_1 dt \qquad t \subset [t_1, t_2], \qquad (1\text{-}3)$$

where

$$A_1 = F_o + \psi(t_2) - \psi(t_1) \, , \qquad B_1 = f_o - (\partial\psi/\partial x)f - (\partial\psi/\partial t) \; . \qquad (1\text{-}4)$$

Here $(\partial\psi/\partial x)$ is n-dimensional vector of partial derivatives. The global minimum is

$$A_0 = \inf_{x_1, x} A_1(x_1, x_2), \quad B_0 = \inf_{v,x} B_1(t,x,v) \quad for \forall t \in T \; . \qquad (1\text{-}5)$$

Depending on the nature of the functions used for ψ, a different set of algorithms for obtaining the infimum can be developed. For example, if Eq.(1-2) takes the form

$$\psi = \lambda(t)x \; , \qquad (1\text{-}6)$$

where $\lambda(t)$ is an n-dimensional vector, the global minimum functions can be written as,

$$A_0 = \inf_{x_1, x_2} [F_0 + \psi(t_2) - \psi(r_1)], \quad B_0 = \inf_{x,v} [\, f_0 - \lambda f(t,x,v) = d\psi / dt] \equiv \inf_{x,v} B \qquad (1\text{-}7)$$

Using $\partial B/\partial x = 0$ and Eq.(1-7) gives

$$\frac{d\lambda}{dt} = -\frac{\partial H}{\partial x}, \quad \overline{B} = \inf_{v \in V} B \quad , \qquad (1\text{-}8)$$

where $H = \lambda f(t,x,u) - f_o$.

Eq. (1-8) can be integrated to find ψ, to obtain the optimal control v and the optimal trajectory $x(t)$. Another way is to enforce the condition

$$B = \inf_{v \in V} \left[f_0 - \frac{\partial \psi}{\partial x} f - \frac{\partial \psi}{\partial t} \right] = 0 \, , \qquad (1\text{-}9)$$

everywhere in the admissible domain for *x*. In this case, the equation for particular derivatives can be solved and the syntheses of the optimal control $v = v(t,x)$ and the field of the optimal trajectory in the admissible domain is obtained.

The two control design approaches with constraints on the maximum control forces are discussed in this section. In the first section an objective function for establishing of the minimum time to suppress vibration is discussed and in the second, the quadratic function is minimized.

A. Minimum Time

Since the main purpose of the controller is to suppress vibrations in minimum time, the time for the system to come to rest is taken as the objective function. A functional expression for this can be written

$$I = \int_0^T dt, \qquad T = min \tag{1-10}$$

subject to

$$dx/dt = Ax + bf, \quad x(0) = x_o, \quad x(T) = 0 \tag{1-11}$$

with control force limits

$$|f_i| \leq F_i \ , \quad i = 1,2,\ldots,p \ . \tag{1-12}$$

This problem can be written in short form as

$$\min I = \int_0^T dt, \qquad dx/dt = Ax + bf, \ x(0) = x_o, \ x(T) = 0, \ |f| \leq F, \tag{1-13}$$

where x is the state vector of dimension $2n$. A is the $2n\times 2n$ plant matrix, B is $2n\times p$ control matrix, f is the control force vector of dimension p, $x(0)$ is the initial state vector, and $x(T) = 0$ is the final state of the system. B_o, in Eq.(1-7), for this problem can be written as

$$B_o = 1 - \Sigma(\partial\psi/\partial x_i)(dx_i/dt) - (\partial\psi/\partial t) \quad (i = 1.,,, n). \tag{1-14}$$

Substituting

$$\psi = \Sigma\lambda_i(t)x_i \qquad (i = 1,2,\ldots,2n) \tag{1-15}$$

and Eq.(1-11) into Eq. (1-14) gives

$$B = 1 - \sum_{J=1}^{2n} \lambda_J \left(\sum_{i=1}^{2n} a_{ij} + \sum_{k=1}^{p} b_{jk} f_k \right) - \sum_{j=1}^{2n} \dot{\lambda}_j x_j \quad . \tag{1-16}$$

Taking the partial derivatives of B $(\partial B/\partial x_i)$ gives

$$d\lambda_j/dt = -\sum_j a_{ij}\lambda_j \qquad i = 1,2,\ldots,n, \quad j = 1, 2,\ldots,2n \quad . \qquad (1\text{-}17)$$

min B_o gives the control force f

$$f_i = |F_i| \,\mathrm{sign}\,(\Sigma b_{jk}\,\lambda_k) \qquad i = 1, 2,\ldots,p\,; \; k = 1, 2,\ldots,2n\,. \qquad (1\text{-}18)$$

Using Eqs.(1-13),(1-17) and (1-18), the optimal control force $f_i(t)$ and trajectory $x_i(t)$ can be calculated. However the initial $\lambda_i(0)$ for our trajectory with $x(0) = x_o$ is not known. To find $\lambda_i(0)$, any suitable gradient method can be used. For example, if the assume some initial state $\lambda_i(0)$ and integrate Eqs.(1-13),(1-17) and (18), we can calculate the function

$$I = T + \sum_{i=1}^{2n} C_i x_i^2(T), \qquad C_i > 0 \;, \qquad (i=1, 2, \ldots,2n). \qquad (1\text{-}19)$$

Here C_i are weight coefficients. If $\sum C_i x^2{}_i(T) < C_0$, where C_o is small, the problem can be considered as solved. Time is optimal and $x_i(t)$ is the optimal trajectory which satisfies the final condition $x_i(T) = 0$. If $\sum C_i x^2{}_i(T) > C_o$ we can choose a new $\lambda i(0)$ by any method and repeat the process until it satisfies $\sum C_i x^2{}_i(T) < C_o$.

In practice, a new independent variable τ is introduced as $t = c\tau$; which can be included with Eq.(1-11) to prouder the additional equation

$$dt/d\tau = c \;. \qquad (1\text{-}20)$$

Additionally, introducing a fixed interval of integration $[0, \tau_1]$ a new set of equations become

$$\min I = \int_0^{\tau} c\,d\tau\,, \quad dx/dt = (Ax + Bf)c\,. \;\; x(o) = x_o, \;\; x(\tau_1) = 0, \;\; |f| \le F \;, \qquad (1\text{-}21)$$

where c is some constant, which is also selected. Eq.(1-19) thus becomes

$$I_1 = \Sigma\, C_i x^2{}_i(\tau_1)\,, \quad i = 1,2,\ldots,2n \quad . \qquad (1\text{-}22)$$

For the structural system as defined by Eqs.(1-11)-(1-12) this problem can be solved for the case in which the number of control inputs, p, is equal to the number of modeled structural degrees of freedom, n. However, numerical difficulties would be encountered when this condition is not satisfied. Typical difficulties would be the occurrence of many local minimums, poor convergence, and the need for smaller step sizes.

B. Linear quadratic regulator (LQR) with bounded control

In this case, a performance index, J, is defined as

$$J = \int_0^{\infty} (x^T \underline{Q} x + f^T \underline{R} f)\,dt \qquad t \in [0,\infty] \quad . \qquad (1\text{-}23)$$

Where $\underline{Q}$ and $\underline{R}$ are state and control weighting matrices. The matrix $\underline{Q}$ must be positive semi definite ($x^T\underline{Q}x \geq 0$), and $\underline{R}$ must be positive definite ($f^T\underline{R}f > 0$). The dimensions of $\underline{Q}$ and $\underline{R}$ depend on the size of vectors the x and f, respectively. The matrices $\underline{Q}$ and $\underline{R}$ can be written as

$$\underline{Q} = \delta Q \tag{1-24}$$

and

$$\underline{R} = (1/\gamma)R^{-1} \tag{1-25}$$

where σ and γ are the design positive variables and $\underline{Q}$ and R^{-1} are constant identity matrices.

The weighting matrix R is defined in terms of the inverse of the constant matrix R in order to maintain positive definiteness. The function B, Eq. (1-9) for the performance index defined in Eq.(1-23) and the constraint equation Eq.(1-11), become

$$\overline{B} = \inf_f \left[(\delta x^T Qx + \gamma f^T Rf) - \frac{\partial \psi}{\partial x}(Ax + Bf) - \frac{\partial \psi}{\partial t} \right] = 0 \tag{1-26}$$

If V represents an open domain, the function ψ, can be written in the form

$$\psi = x^T Px\,, \tag{1-27}$$

where P is a $2n$-dimensional unknown matrix.

Substituting ψ Eq. (1-27) into Eq, (1-26), we obtain the equation

$$\sigma\gamma Q + PA + A^T P - \gamma PBRB^T P = 0 \quad . \tag{1-28}$$

Equation (1-28) is the Riccati equation. A solution of this equation gives the matrix P and one can find the optimal control force as

$$f = -Gx \tag{1-29}$$

where

$$G = \gamma RB^T P \tag{1-30}$$

Integrating Eq. (1-11) using Eq. (29) to obtain the optimal trajectory for the LQR functional. Eq. (1-29) may give unrealistic values of control depending on the selection of γ. The magnitude of control can be decreased by increasing γ, however, this may cause other perturbations of the system (such as the time it takes the oscillation to decay) to deteriorate.

In order to obtain more realistic results, bounds can be placed on the control force. This can be written as

$$|f_i| \leq F_i\,, \quad F_i = \text{const}, \;\; i = 1, 2, \ldots, p \tag{1-31}$$

where F_i is the magnitude bounding each controller. To obtain an optimal solution, the following restrictions must be satisfied: (1) among these optimal synthesis of the control must exist in the domain of interest, (2) the function B Eq. (1-30) must be convex, and (3) the limits of F may be constant or dependent on time only and F must not be equal to zero at any time (Note: if F is very small a loss in stability can occur). For a solution, the system of Eqs. (1-11) and (1-29) must be integrated along with limits imposed by equation (1-31).

The norm for the displacements or total deviation can be defined by

$$R_x(t) = S = [\sum_{i=1}^{n} x_i^2(t)]^{1/2} \quad i = 1, 2, \ldots, n \qquad (1\text{-}32)$$

This norm is zero at the time the deviation is zero, and the structure stops vibrating. In the LQR solution domain this time equals infinity. For studying the behavior and comparison of different control systems, a measure of performances has been used based upon. The time required to reduce the norm of the displacements to 2% of their initial value.

Numerical Examples.

Example 1. SISO problem.

For comparison of systems with different objective functions, a vibrating structure with a single physical degree of freedom was been investigated. This system is described by equation the following set of

$$dx_1/dt = x_2, \quad dx_2/dt = -\omega^2 x_1 - 2\zeta\omega x_2 + cf, \;\; x_1(0) = 0, \;\; x_2(0) = 1, \;\; |f| \le 1 \qquad (1\text{-}33)$$

where $\omega = 2$ is the frequency, $\zeta = 0.03$ is the damping, $c = 1$, and $|f| \le 1$ is the control.

The problem is solved having an objective function for minimum time as

$$\min T = \int_0^T dt, \;\; x_1(T) = 0, \;\; x_2(T) = 0 \qquad (1\text{-}34)$$

Eqs. (1-17) and (1-18) for the system defined in Eq. (1-33) become

$$d\lambda_1/dt = -\omega^2\lambda_2, \quad d\lambda_2/dt = \lambda_1 - 2\xi\omega\lambda_2, \quad f = |F| \operatorname{sign} \lambda_2 . \qquad (1\text{-}35)$$

Eqs. (1-33)-(1-35) are integrated and the initial values $\lambda_1(0)$, $\lambda_2(0)$ are chosen such that the conditions $x_1(T) = x_2(T) = 0$ are satisfied. The details of the solution scheme are not given here because of space limitations.

The performance for the linear quadratic regulator (LQR) is

$$J = \int_0^\infty \frac{1}{2}\left[(\delta_1 x_1^2 + \delta_2 x_2^2\right]dt \;. \qquad (1\text{-}36)$$

Using this performance index and solving the Riccati Eq. (1-29) gives

$$f = 2(c/\gamma)(c_{12}x_1 + c_2 x_2), \qquad (1\text{-}37)$$

where

$$c_{12} = -[\omega^2 + (\omega^4 + c_o\delta_1)^{0.5}]/2c_o, \;\; c^2 = \{-\zeta\omega + [\xi^2\omega^2 + (0.25\delta_2 + c_{12})c_o]^{0.5}\}/c_o, \quad c_o = c^2/\gamma .$$

In the case of $\delta = \delta_1 = \delta_2$, the time history depends only on δ/γ. The total deviation is

$$R_x = S = (x_1^2 + x_2^2)^{1/2} \;. \qquad (1\text{-}38)$$

Eq. (1-33) is integrated with control given in Eq. (1-37).

The results of this investigation for the case T = min, σ/γ = 0.25 and 100 and no control (open-loop system) are shown in Fig's 1, 2, & 3.

Fig. 1 shows the time history of deviation of x_2. As can be seen, an LQR with σ/γ = 100 gives better results (t = 4 sec) than an LQR with σ/γ = 0.25 (time is more than 15 sec) however an even better result is obtained with an objective function of minimum time. In the last case, oscillations are terminated in 1.5 sec.

Figure 2 shows the variation of a bounded control force $|f| \leq 1$ for the case of T=min, LQR when σ/γ = 0.25 and σ/γ = 100. The case LQR (σ/γ = 0.25) does not use the full control force, the case LQR (σ/γ =100) uses more of the control force, and case t =min uses the maximum control force all the time.

Fig.3 shows the time history for the total deviation (R_x) with no control, with an objective function for minimum time and with LQR given by control bounds $|f| \leq 1$.

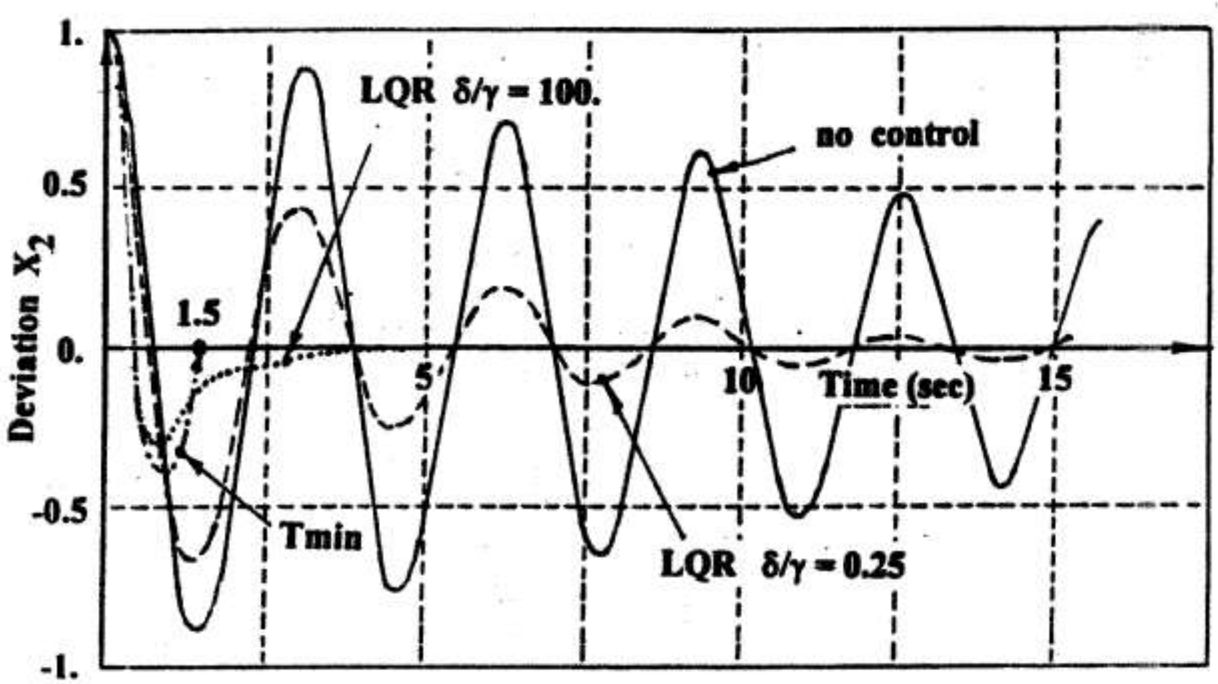


Fig 1

A structural system with any number of degrees of freedom can be transformed into pairs of equations (1-33)(see later Eq.(1-40)-(1-48)) where every pair is independent from the other. If the number of controls equals the number of degrees of freedom the design approach based on minimum time can be used. However, if the number of controls is less than the number of pairs of equations, the solution for the functional T = min becomes very complex. In this case, the LQR approach is a variable alternative.

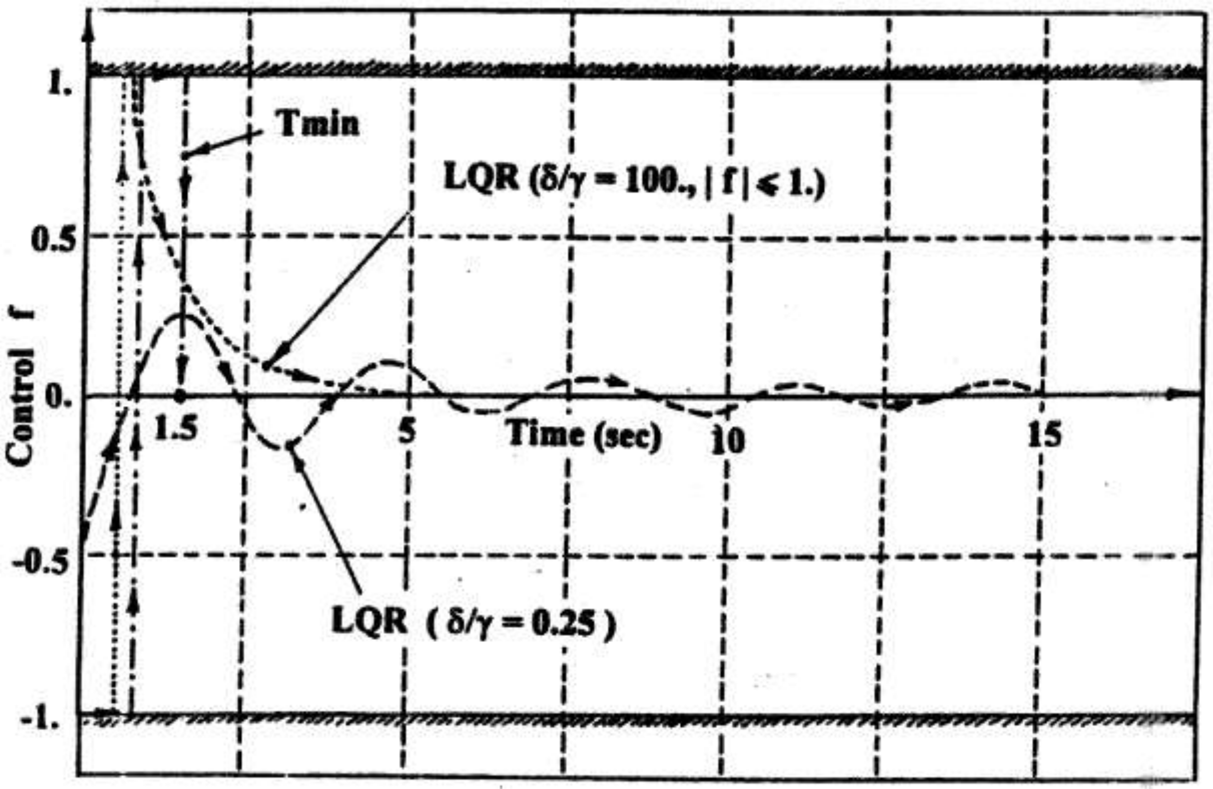


Fig.2

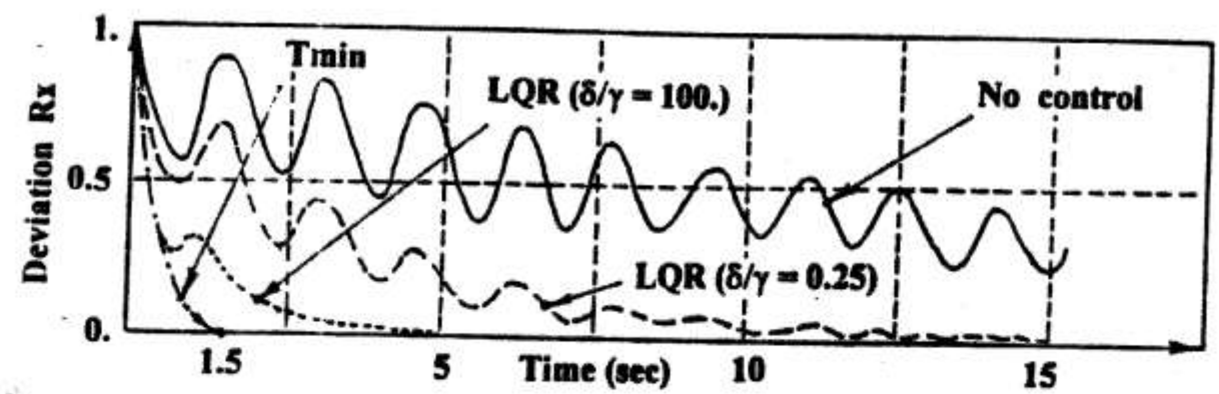


Fig.3

Example 2. Wing Box

In order to illustrate the application of an approach using the linear quadratic regulator with bounded control, the wing box problem in reference [5] is used and shown in Fig. 4. This structure has thirty-two elements and twenty-four degrees of freedom. The structure is a cantilever wing box idealized with bar elements capable of carrying axial loads only.

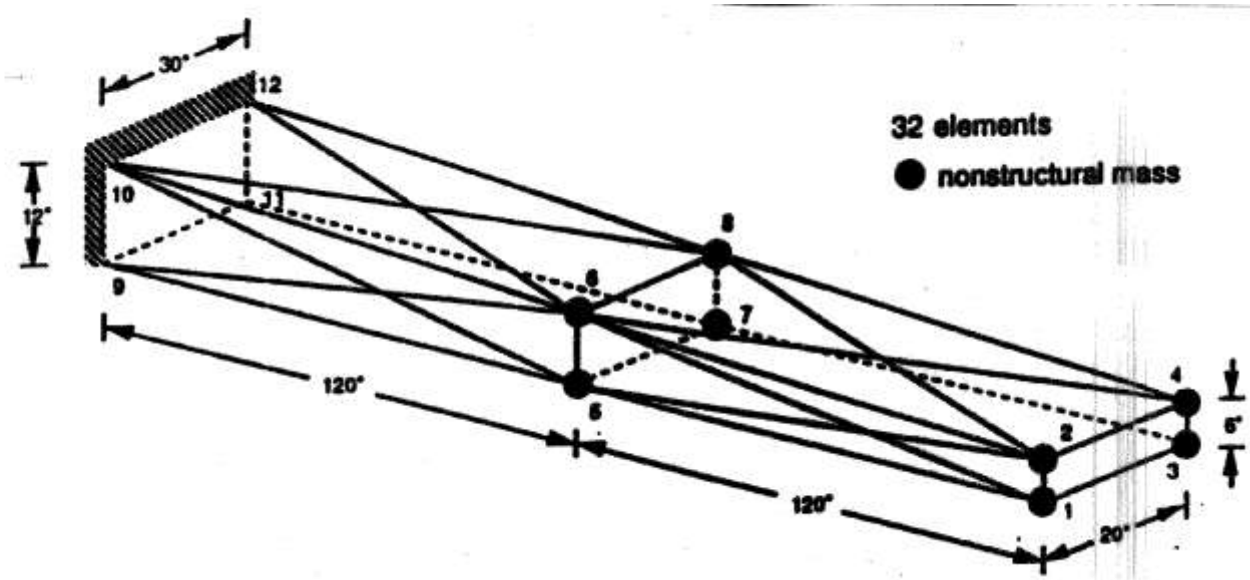


Fig.4.

The equations of motion for a flexible structure with no external disturbance can be written as

$$M\ddot{u} + E\dot{u} - Ku = Df \ , \qquad (1\text{-}39)$$

where M is the mass matrix, E is the damping matrix, and K is the total stiffness matrix. These matrices are $n_1 \times n_1$, where n_1 is the number of degrees of freedom of the structure. In Eq. (1-39), D is the applied load distribution matrix relating the control input vector f to the coordinate system. The number of elements in vector f is equal to the number of actuators, p. The vector u in Eq.(1-39) defines the structural response.

The coordinate transformation

$$u = [\Phi]\eta \qquad (1\text{-}40)$$

is introduced where η is the modal coordinate system and [Φ] is the $n_1 \times n_1$ modal matrix. Using Eq. (1-40), Eq.(1-39) can be transformed into n_1 uncoupled equations. These can be written as

$$\overline{M}\ddot{\eta} + \overline{E}\dot{\eta} + \overline{K}\eta = [\Phi]^T Df \qquad (1\text{-}41)$$

where

$$\begin{aligned} \overline{M} &= I = [\Phi]^T M[\Phi] \\ \overline{E} &= [2\zeta\omega] = [\Phi]^T E[\Phi] \\ \overline{K} &= [\omega^2] = [\Phi]^T K[\Phi] \end{aligned} \qquad (1\text{-}42,\ 1\text{-}44)$$

The matrices $\overline{M}, \overline{E},$ and $\overline{K}$ are diagonal square matrices, ω is the vector of structural frequencies, and ζ is the vector of modal damping factors. The modal matrix [Φ] is normalized with respect to the mass matrix. The control analysis is performed by reducing the second-order uncoupled equation [Eq.(1-41)] to a first-order equation. Only n of n_1 uncoupled equations are used for the control system design. This can be achieved by using the transformation

$$x_{2n} = \begin{bmatrix} \eta \\ \dot{\eta} \end{bmatrix}_{2n} \tag{1-45}$$

where x is the state variable vector of size $2n$. This gives

$$\dot{x} = Ax + Bf \tag{1-46}$$

where A is a $2n \times 2n$ matrix and B is a $2n \times p$ input matrix. The A matrix and the input matrices are given by

$$A = \begin{bmatrix} 0 & I \\ -\omega^2 & -2\xi\omega \end{bmatrix} \tag{1-47}$$

$$B = \begin{bmatrix} 0 \\ \Phi^T D \end{bmatrix} \tag{1-48}$$

The state output equation is given by

$$y = Cx \tag{1-49}$$

where y is a $q \times 1$ output vector, C is a$q \times 2n$ output matrix, and q is equal to the total number of sensors. If the number of sensors and actuators equal and collocated, then $q = p$ and

$$C = B^T . \tag{1-50}$$

For this structure, Young's modulus and weight density are assumed to be equal to 10.5 x $l0^6$ lbs/in^2 and 0.1 lbs/in^3, respectively. The actuators and sensors are assumed embedded in the structural elements and are collocated. The actuators are assumed to apply forces along the axial directions providing both out of plane, in plane and twist control for the structure. It is assumed that all structural modes have 1% structural damping and thus ζ in Eq. (1-9) was 0.01.

The control system utilizes four actuators and sensors collocated in the four members at the tip of the structure connecting nodes 1-2, 3-4, 1-3 and 2-4 respectively. Non-structural masses are located at nodes 1 through 8. Their magnitudes are 0.5 slugs at nodes 1 and 2; 1.5 slugs at nodes 3 and 4; 2.5 slugs at node 5 and 7.0 and 1.0 slugs at nodes 6 and 8 respectively. For the 24 structural degrees of freedom, the full order state space matrix in Eq. (1-11) is 48 x 48. Since there are four actuators and sensors, the input matrix B and output matrix C are 48 x 4 and 4x48, respectively. The cross-sectional areas of the rod elements were equal to 0.1 in^2. The weighting matrices Q and R in Eq. (1-28), (1-29) were equal to the identity matrix.

The four values of the weighting parameter ratios σ/γ selected for this study are 0.1, 1.0, 100 and 1000, respectively. The maximum control forces generated by the four actuators are given in Table **1.**

Table 1. Calculated cases

Control bound F	∞	0.5	0.15	0.05

δ⁄γ = 0.1	+	+	+	+
δ⁄γ = 1	+	+	+	+
δ⁄γ = 100	+	+	+	+
δ⁄γ = 1000	+	+	+	+
No control	+			

The initial condition used for designing the controllers is a unit displacement at node 1 in the z-direction. This condition is used for all cases and also to obtain the response curves. The response curves are given for only a few cases because of space limitations. The three limits on the maximum allowable control forces are set equal to 0.5, 0.15, and 0.05 respectively. The different cases considered are summarized in Table 2.

Table 2**. Maximum actuator forces**

	Actuator #			
Value δ⁄γ	1	2	3	4
δ⁄γ = 0.1	0.05	0.05	0.07	0.03
δ⁄γ = 1.0	0.20	0.24	0.31	0.12
δ⁄γ = 100	1.31	1.89	3.30	1.23
δ⁄γ = 1000	2.95	2.25	8.25	3.73

In the case of σ⁄γ = 0.1 the maximum actuator forces are less than 0.15, and for σ⁄γ = 1.0, they are less than 0.5. Fig. 5 shows the time history of the displacement norm without control bound for the four values of σ⁄γ and without control.

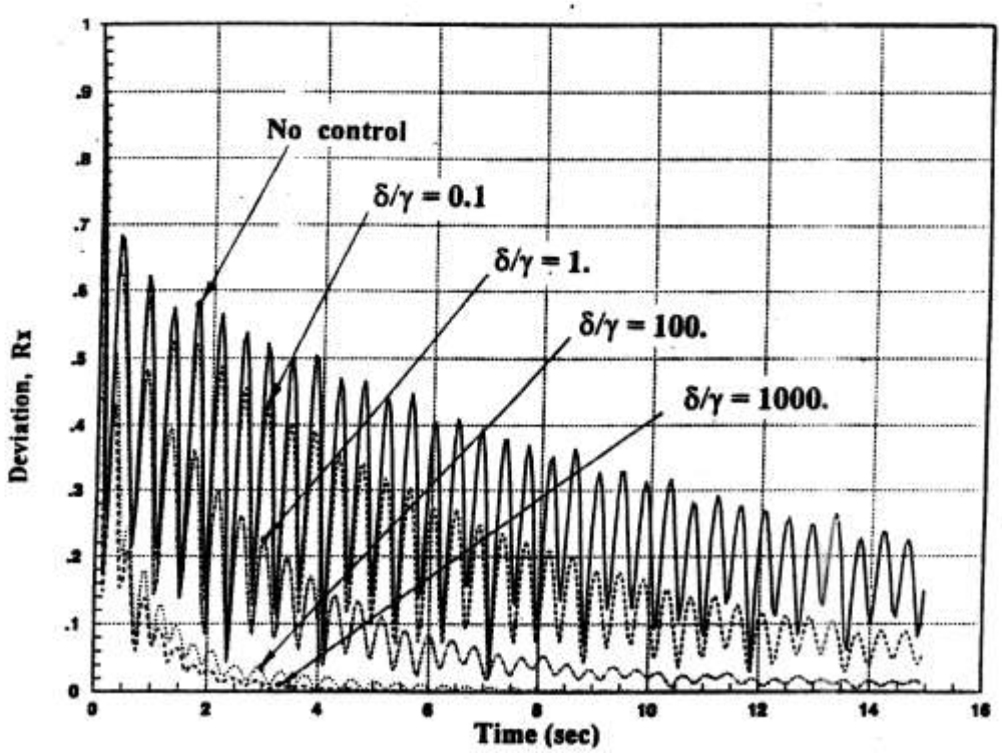
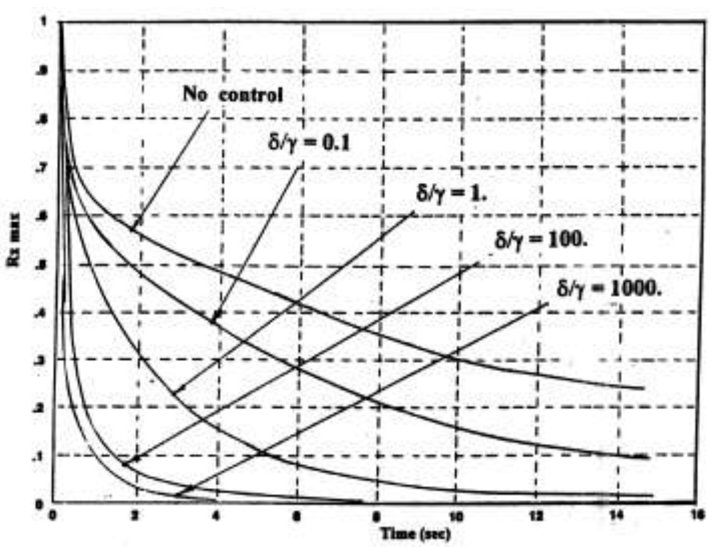


Fig.5.

The maximum value of the displacement norm as a function of time is shown in Fig. 6.

The time required to decrease the displacement norm to 2% of its initial value 1.0 is shown in Fig. 7.



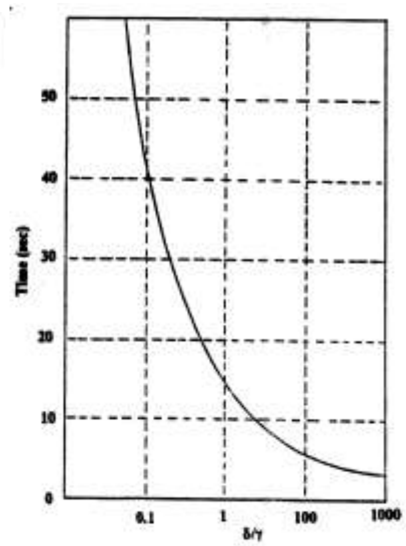


Fig.6.

Fig.7.

In the case of no-control, the total time needed to reduce the displacement norm to two percent of the initial value is larger than 100 seconds. The variation in the control force in actuator 1 as a function of time for σ/γ equal to 100 and 1 is shown in Fig. 8.

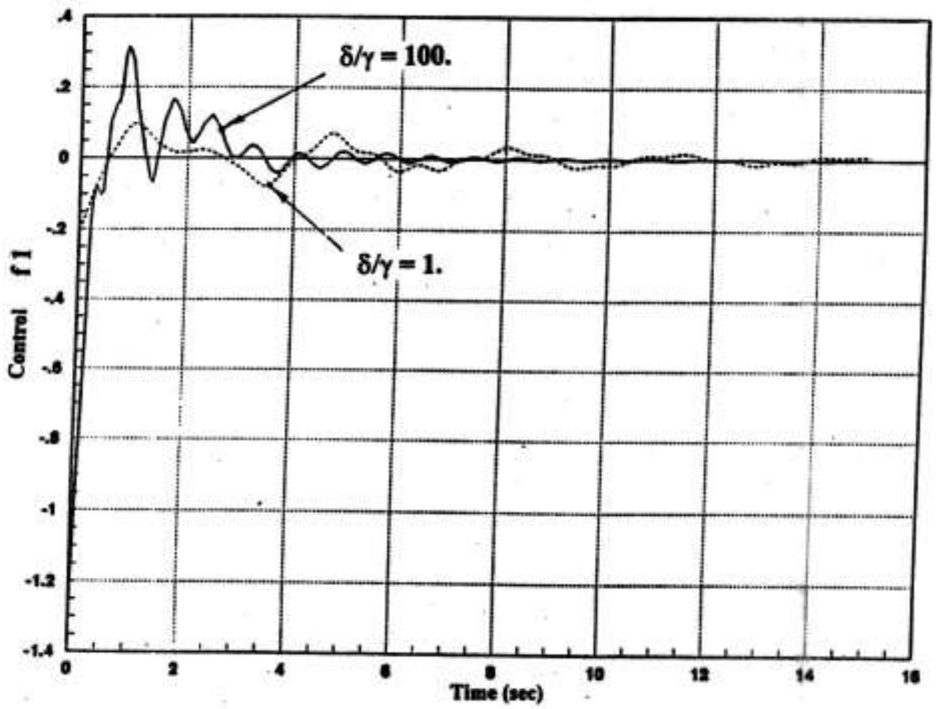


Fig.8. (2.47 Mb)

Fig. 9 shows the time history of control force in actuator 1 with the upper bound equal to 0.15 for σ/γ = 100. The upper bound is enforced on all the actuators.

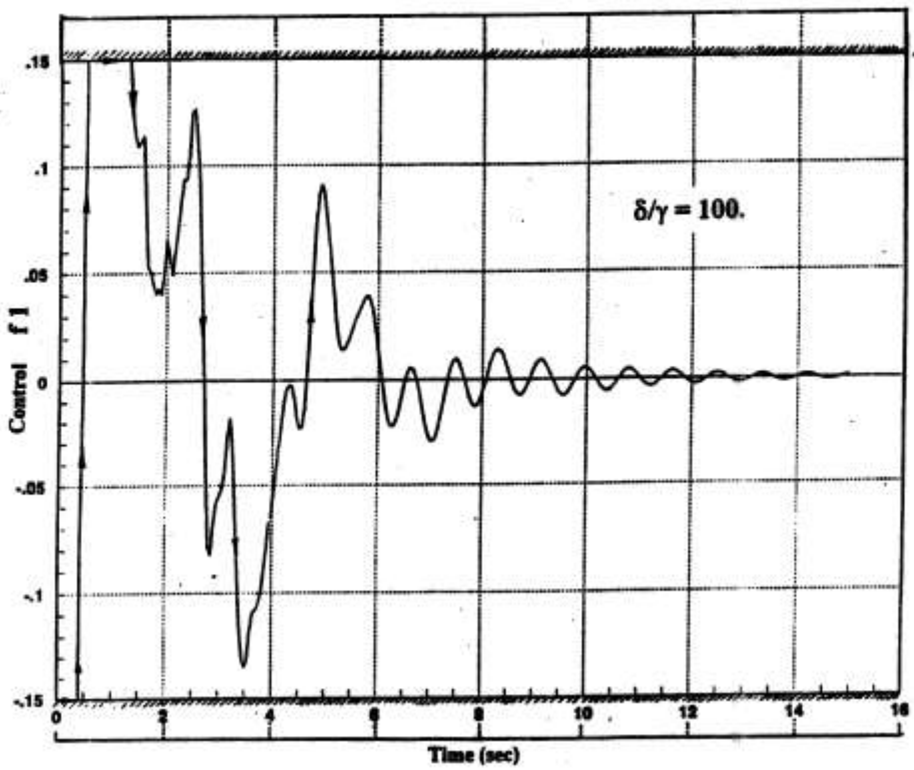


Fig.9.

The changes in the displacement norm with time for 8/y equal to 100 are shown in Fig. 10 for the case of control bound equal to 0.15 and without bound.

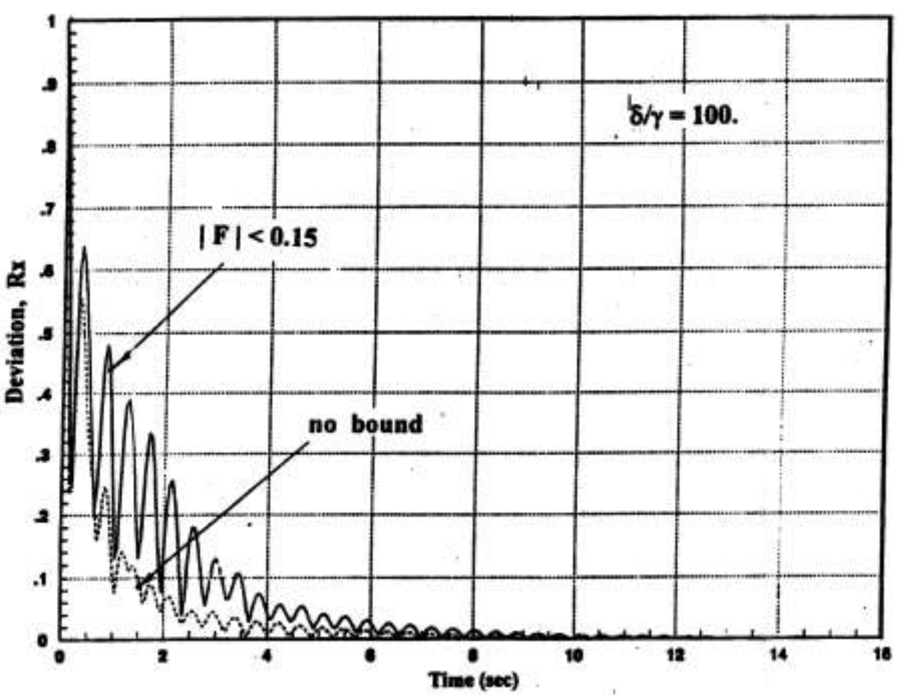


Fig.10.

Fig. 11 shows the total time required to reduce the displacement norm to 0.02 for three values of σ/γ and four values of control bound. As the control bound decreases more time is needed to reduce the displacement norm to 0.02 for a given value of σ/γ. The maximum root mean square response for different cases is shown in Fig. 12.

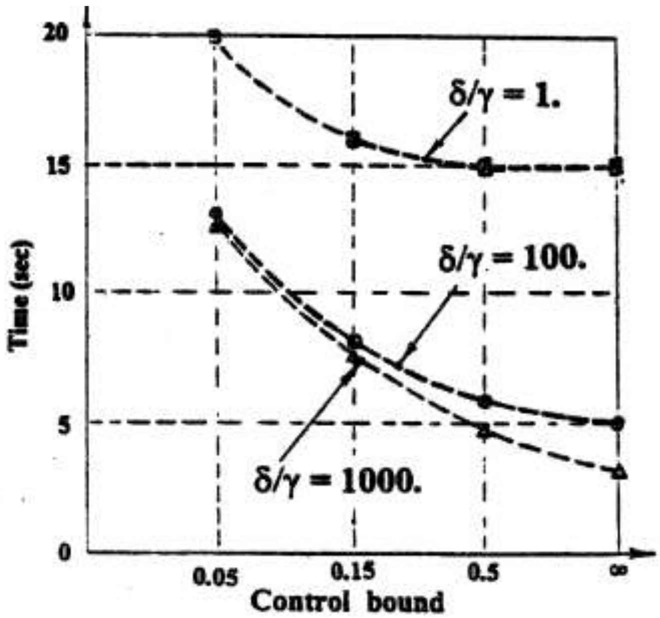


Fig.11.

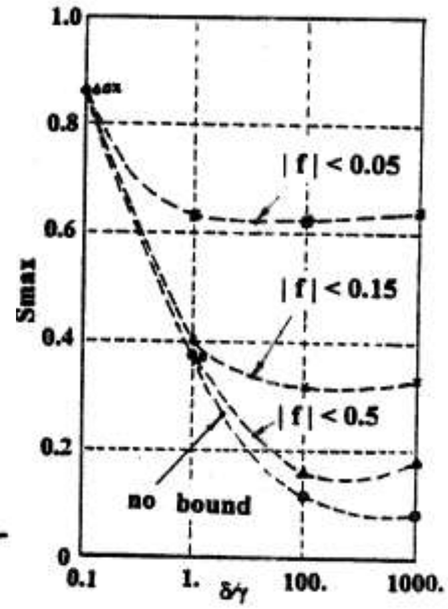


Fig.12

2. Solution of general linear optimal problem for one control

Now consider the general optimal linear regulator problem with an objective function of minimum time and one control parameter.

<u>Problem Statement</u>. The system is described by a linear differential equation in vector form as,

$$\dot{x} = Ax + Lu \qquad (2\text{-}1)$$

where $x = (x_1, x_2,..., x_n)$ is the *n*-dimensional state vector , $A = \|a_{ij}\|$ a *n*-dimensional square matrix of constant coefficients, *L* a column vector which contains $l_1, l_2,..., l_n$; *u* a limited control, $|u| \le \zeta$, $\zeta > 0$; $x(0) = x_o$, $x(t_k) = x_k$ the initial and final condition, $T = t_k$ represent the end time of process, $t_o = 0$.

It is known the control can have only boundary value in linear system and, if eigenvalues of matrix *A* is real numbers, the system has only maximum *n-1* switches [7].

<u>Problem</u> <u>solution</u>. The characteristic equation is $|A - \lambda E| = 0$, where *E* is an unit *n*-dimensional matrix, λ is eigenvalues of matrix *A*.

<u>Case A</u>. All eigenvalues λ are real, different, and not equal zero. Using

$$y_i = \sum_{j=1}^{n} e_{ij} x_j \qquad (i=1,2,\ldots,n), \qquad e_{ij} = const_{ij}$$

can convert the equations (2-1) to canonical form

$$\dot{y}_1 = \lambda_1 y_1 + b_1 u; \quad \dot{y}_2 = \lambda_2 y_2 + b_2 u; \quad \ldots \quad \dot{y}_n = \lambda_n y_n + b_n u; \qquad (2\text{-}2)$$

with boundary conditions $y_i(0) = y_{i0}$; $y_i(t_k) = y_{ik}$.

The optimal control $u = \pm\zeta$ is constant everywhere. If a new variable

$z_i = \lambda_i y_i + b_i u$ is introduced, it is possible to write equation (2-2) in form

$$\dot{z}_1 = \lambda_1 z_1; \quad \dot{z}_2 = \lambda_2 z_2; \quad \ldots; \quad \dot{z}_n = \lambda_n z_n; \qquad (2\text{-}3)$$

A solution of equation (2-3) is

$$z_i = \bar{c}_i e^{\lambda_i t} \quad (i = 1, \ldots, n).$$

Returning to the variable y we can write

$$y_i = c_i e^{\lambda_i t} - b_i u / \lambda_i \quad (i = 1,\ldots,n); \qquad c_i = \bar{c}_i / \lambda_i; \quad \lambda_i \neq 0. \qquad (2\text{-}4)$$

Consider the value y_1. The moment when a control parameter is changed it is marked an index "i" below and right and left from point t_i by plus and minus sign on top of magnitudes.

Let us suggest, that the control has $k-1$ switches. From continuous condition we have

$y_i^- = y_i^+$. Therefore we have

$$c_{ji}^- e^{\lambda_{ji} t_i} - b_i u_i^- / \lambda_i = c_{ji}^+ e^{\lambda_{ji} t_i} - b_i u_i^+ / \lambda_i \qquad (i = 1,\ldots,n) . \qquad (2\text{-}5)$$

From (2-5)

$$c_{ji}^+ = c_{ji}^- + e^{-\lambda_j t_i} (u_i^+ - u_i^-) b_j / \lambda_i \qquad (i = 1,\ldots,k\text{-}1). \qquad (2\text{-}6)$$

The value $c_{j,i+1}^- = c_{j,i}^+$. From (2-6), we get

$$c_{jk}^- = c_{j0}^+ + \sum_{i=1}^{k-1} \frac{b_j}{\lambda_i} e^{-\lambda_i t_i} (u_i^+ - u_i^-) \ . \qquad (2\text{-}7)$$

From the first equation (2-4) and boundary conditions for y_i, we find

$$c_{jk}^{-} = e^{-\lambda_j t_k t_k}(y_{jk} + \frac{b_j}{\lambda_j}u_k^{-}); \qquad c_{j0}^{+} = y_{j0} + \frac{b_j}{\lambda_i}u_0, \qquad (2\text{-}8)$$

where $u_0 = u_0^{+}$.

Substituting (2-8) to (2-7) we obtain,

$$\sum_{i=1}^{k-1}\frac{b_j}{\lambda_j}e^{-\lambda_i t_i}(u_i^{+} - u_i^{-}) - e^{-\lambda_j t_k t_k}(y_{jk} + \frac{b_j}{\lambda_j}u_k^{-}) = -y_{j0} + \frac{b_j}{\lambda_i}u_0 \qquad (j=1,\dots,n) \qquad (2\text{-}9)$$

These equation (2-9) satisfies for all y_i i=1,…,n.

If to divide the right and left parts of equation (2-9) by $(-2b_j u_o/\lambda_l)$, we find,

$$e^{-\lambda_1 t_1} - e^{-\lambda_1 t_2} + \dots - (-1)^k e^{-\lambda_1 t_{k-1}} + \frac{1}{2}\left[(-1)^{k-1} + y_{1k}\lambda_1 / b_1 u_0\right]e^{-\lambda_1 t_k} = \frac{1}{2}\left[1 + y_{10}\lambda_i / b_1 u_0\right];$$

$$e^{-\lambda_2 t_1} - e^{-\lambda_2 t_2} + \dots - (-1)^k e^{-\lambda_2 t_{k-1}} + \frac{1}{2}\left[(-1)^{k-1} + y_{2k}\lambda_2 / b_2 u_0\right]e^{-\lambda_2 t_k} = \frac{1}{2}\left[1 + y_{20}\lambda_2 / b_2 u_0\right];$$

. . . . (2-10)

$$e^{-\lambda_n t_1} - e^{-\lambda_n t_1} + \dots - (-1)^k e^{-\lambda_n t_{k-1}} + \frac{1}{2}\left[(-1)^{k-1} + y_{nk}\lambda_n / b_n u_0\right]e^{-\lambda_n t_k} = \frac{1}{2}\left[1 + y_{n0}\lambda_n / b_n u_0\right],$$

where k=n.

Noting that $e^{-t_i} = w_i$, equations (2-10) can be written as

$$w_1^{\lambda_1} - w_2^{\lambda_1} + \dots - (-1)^k w_{k-1}^{\lambda_1} + \frac{1}{2}\left[(-1)^{k-1} + y_{1k}\lambda_1 / b_1 u_0\right]w_k^{\lambda_1} = \frac{1}{2}\left[1 + y_{10}\lambda_1 / b_1 u_0\right],$$

$$w_1^{\lambda_2} - w_2^{\lambda_2} + \dots - (-1)^k w_{k-1}^{\lambda_2} + \frac{1}{2}\left[(-1)^{k-1} + y_{2k}\lambda_2 / b_2 u_0\right]w_k^{\lambda_2} = \frac{1}{2}\left[1 + y_{20}\lambda_2 / b_2 u_0\right],$$

. . . . (2-11)

$$w_1^{\lambda_n} - w_2^{\lambda_n} + \dots - (-1)^k w_{k-1}^{\lambda_n} + \frac{1}{2}\left[(-1)^{k-1} + y_{nk}\lambda_n / b_n u_0\right]w_2^{\lambda_n} = \frac{1}{2}\left[1 + y_{n0}\lambda_n / b_n u_0\right].$$

Equation (2-11) is solved in order to find $w_i = w_i(y_o)$. Returning to the original variable x, we can write

$$t_i = -\ln w_i(x).$$

where x_o represents the initial point x.

Equations (2-11) are a set of algebraic equations. From boundary conditions we know that

$$t_k \geq t_{k-1} \geq \ldots \geq t_1 > 0 \ .$$

That implies that

$$0 < w_k \leq w_{k-1} \leq \ldots \leq w_1 \leq 1 \ .$$

For control $u = \pm\xi$. This implies that equations (2-11) must be solved twice. If $x(t_k)$ = 0 (this means $y(t_k)$ = 0), the second solution is symmetric about the origin.

The solution of equation (2-11) is easier to evaluate then the classical optimal control solution. In classical theory a researcher must solve a boundary problem for a set of given differential equations and also find a set of unknown Lagrange multipliers. In using equation (2-10) the researcher first establish the required time increments based upon knowledge of the physical situation.

To find the switch surfaces, for t_1 = 0, implies thus the trajectory is located on the first switch surface. In this case in equation (2-11) $e^{-t_1} = w_1 = 1$. We then set about solving the first *n*-1 equations (2-11) for $w_2, w_3, \ldots, w_k$ and substitute these solutions into the last equation. This leads to an equation $\Phi_1(y) = 0$. By substituting for *y* we can find $N_1(x)$. This is the first (n-1)-dimensional switch surface.

Next by substituting w_1 = 1, w_2 = 1 in the first *n*-1 equations, and solving the first *n*-2 equations for $w_3, w_4, \ldots, w_n$ one can obtain solutions and substitute them into last equation. We thus can find a hyper surface $N_2(x)$ = 0. The intersection of this hyper surface with $N_1(x)$ creates the second (*n*-2)-dimensional switch surface. Other switch surfaces can be found in as similar way.

Such an optimal control result can be easily found. In selecting u_o, when the state point reaches the switch surface $N_1(x)$ = 0, u_1= - u_o. When the state point reaches the switch surface $N_2(x)$ = 0, u_2= -u_1 and so on.

If time is deleted from any two of equation (2-4), we obtain a projection of the trajectory on the surface $y_i y_j$

$$y_i = \bar{c}_i(y_i + b_i u/\lambda_i)^{\lambda_i/\lambda_j} - b_i u/\lambda_i, \qquad \bar{c}_i = c_i / c_j^{\lambda_i/\lambda_j} \ .$$

From (2-2) we can find the boundaries of instability, for positive eigenvalues. For example, if $\lambda_l > 0$, $b_i >$ 0, then $y_i(t_k)$ = 0. The necessary and sufficient condition unstable solution is given by

$$y_i > \frac{b_i}{\lambda_i}\xi \ ; \qquad y_i < -\frac{b_i}{\lambda_i}\xi \, .$$

We have only considered cases when the eigenvalues are real, different, and non-equal to zero. Additional cases have been considered in reference [6] .

Example. Taking any two of equations (2-1) with eigenvalues λ_1, λ_2 ($\lambda_1 \neq \lambda_2, \lambda_1 \neq 0, \lambda_2 \neq 0, \lambda_1 < \lambda_2$), $x(t_k)$ = 0, a canonical form of the equations can be expressed as,

$$\dot{y}_1 = \lambda_1 y_1 + u; \quad \dot{y}_2 = \lambda_2 y_2 + u; \quad y(0) = y; \quad y(t_k) = 0; \quad |u| \leq 1. \qquad \text{(E1)}$$

Equation (2-11) for u_o = 1 can be written as

$$w_1^{\lambda_1} - \frac{1}{2} w_2^{\lambda_1} = \frac{1}{2}\lambda_1 y_1 + \frac{1}{2}; \qquad w_1^{\lambda_2} - \frac{1}{2} w_2^{\lambda_2} = \frac{1}{2}\lambda_2 y_2 + \frac{1}{2}. \qquad \text{(E2)}$$

For w_1 = 1 (simplifying for the case t_1 = 0), w_2 can be obtained from (E2) and thus

$$y_2 - \frac{1}{\lambda_2}\left[1 - (1 - \lambda_1 y_1)^{\lambda_2/\lambda_1}\right] = 0. \qquad \text{(E3)}$$

For u_o = -1 equations (2-11) are

$$w_1^{\lambda_1} - \frac{1}{2} w_2^{\lambda_1} = -\frac{1}{2}\lambda_1 y_1 + \frac{1}{2}; \qquad w_1^{\lambda_2} - \frac{1}{2} w_2^{\lambda_2} = -\frac{1}{2}\lambda_2 y_2 + \frac{1}{2}. \qquad \text{(E4)}$$

Taking w_1 = 1, and using w_2 from (E2), we find,

$$y_2 + \frac{1}{\lambda_2}\left[1 - (1 + \lambda_1 y_1)^{\lambda_2/\lambda_1}\right] = 0. \qquad \text{(E5)}$$

Using a continuity condition $y_1(t_k) = y_2(t_k)$, the relations (E3), (E5) can be written as one relation

$$y_2 - \frac{sign\, y_1}{\lambda_2}\left[1 - (1 - \lambda_1 y_1 sign\, y_1)^{\lambda_1/\lambda_2}\right] = 0. \qquad \text{(E6)}$$

If $y_1 = 0$, $y_2 > 0$, then the relation (E6) is greater then zero. From (E1) we see: y_2 will be decrease faster if u = -1 for $y_2 > 0$ and u = +1 for $y_2 < 0$. This implied that

$$u = -sign\,\lambda_2\; sign\left\{ y_2 - \frac{sign\, y_1}{\lambda_2}\left[1 - (1 - \lambda_1 y_1 sign\, y_1)^{\lambda_1/\lambda_2}\right]\right\}. \qquad \text{(E7)}$$

To find the equations for optimal trajectories. Referring equations (2-4),(2-11) we find

$$y_1 = c_1 e^{\lambda_1 t} - u/\lambda_1; \quad y_2 = c_2 e^{\lambda_2 t} - u/\lambda_2; \quad y_2 = c\,(y_1 + u/\lambda_1)^{\lambda_2/\lambda_1} - u/\lambda_2. \qquad \text{(E8)}$$

The last equation in (E8) gives information in the trajectories as shown in figure 13.

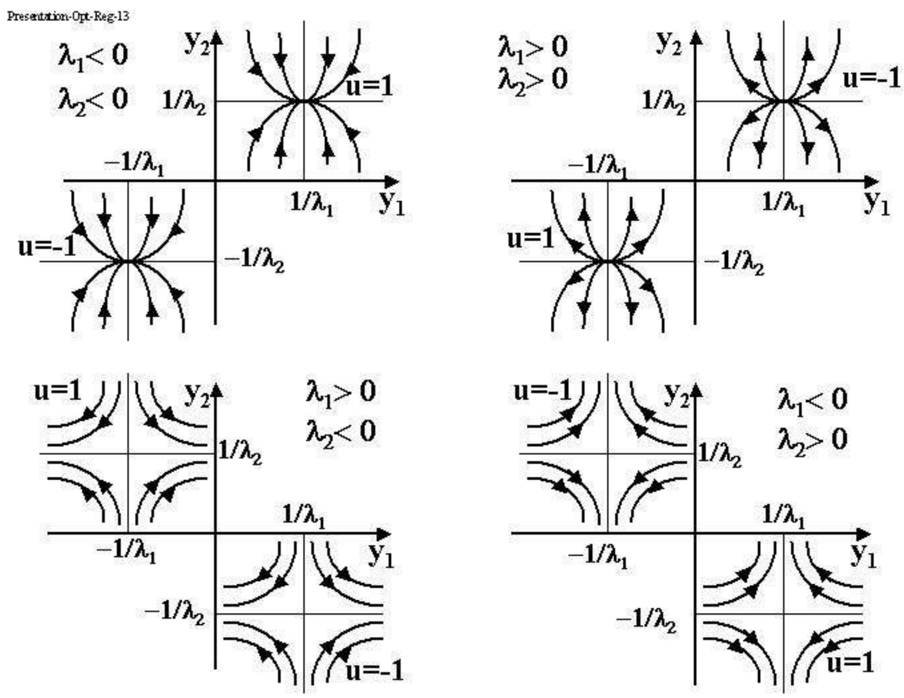


Fig13.

These trajectories depend upon the signs of λ_1, λ_2. For then $\lambda_1>0$, $\lambda_2<0$ the non-stability region is $|y_1| > \zeta/\lambda_1$. For $\lambda_1<0$, $\lambda_2>0$ the non-stability region is $|y_2| > \zeta/\lambda_2$.

In fig.14 also shown optimal trajectories. Once again they depend up on the signs of λ_1, λ_2. Returning to the variables *x*, the picture 14 is affined deformity.

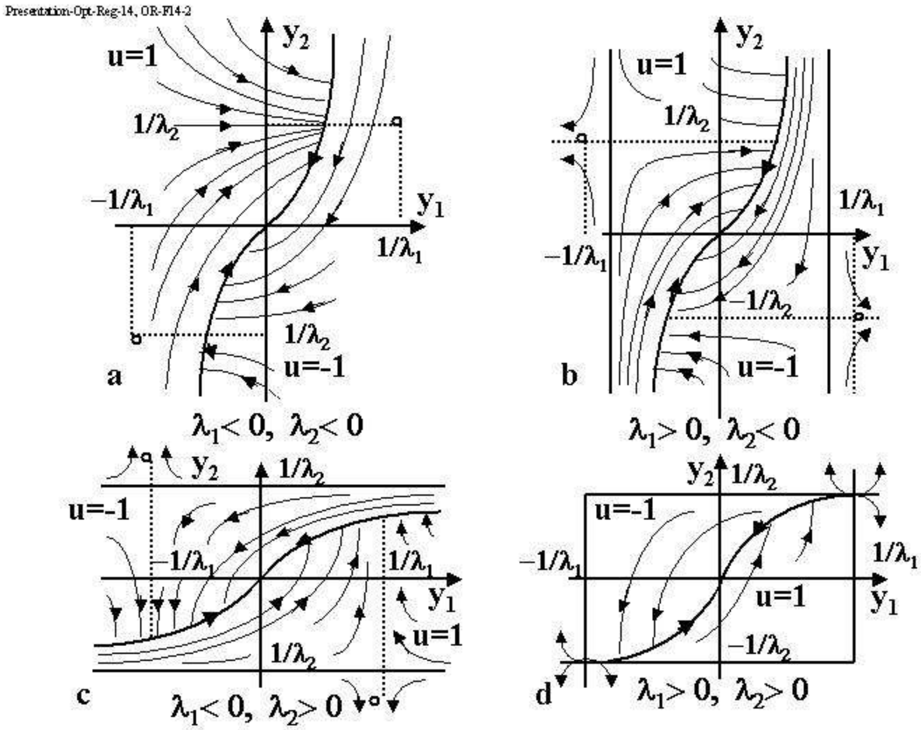


Fig.14.

The offered method allows capture of optimal control.

Summary

Two optimum control design methods for suppression of structural vibration having bounded constrains have been compared. The minimum time and quadratic performance index have been used as objective functions. The second approach leads to use of the LQR methodology with bounded control. The introduction of a minimum time controller can be used when the number of actuators equals the number of structural degrees of freedom used in the design of the control system. When the number of actuators is less than the number of degrees of freedom, the minimum time controller becomes mathematically complicated and has been found to be difficult to solve due to the presence of local minimums. The minimum quadratic function controller, with bounded control, can be designed with a fewer number of actuators. A SISO structural control design problem has been solved using both approaches for comparison of trajectories and the time needed to suppress vibrations. The influence of control limitations and the weight coefficient σ/γ of the structure have been studied. Results indicate that an optimal selection of the weight coefficient σ/γ can decrease the suppression time up to 2-4 times.

Recommendations

If possible, the researcher should try to design the controller for minimum time. If it is very difficult, he can design LQR controller. However in this case the researcher must:

1. Consider limits on the maximum value of the control force.

2. Find the optimal ratio σ/γ of the weight coefficients.
3. Solve (numerically) at least one time the real (minimum time) problem and compare what may be luck is loss from changing the T_{min} problem to the LQR problem.

References

1. A.A. Bolonkin, Solution general linear optimal problem with one control. Journal "Pricladnaya Mechanica", v.4, #4, 1968, pgs. 111-122. Moscow (in Russian).
2. A.A. Bolonkin, New methods of optimization and their application, Moscow, Technical University named Bauman, 1972, pgs.220 (in Russian).
3. A.A. Bolonkin, "Special Extrema in Optimal Control Problems", Akademiya Nauk, Izvestia, Theknicheskaya Kibernetika, No.2, March-April, 1969, pp.187-198. See also English translation in Eng.Cybernetics, n.2, March-April,1969, pp.170-183.
4. A.A. Bolonkin, "A New Approach to Finding a Global Optimum", "New Americans Collected Scientific Reports",Vol.1, 1991. the Bnai Zion Sa Scientists Division, New York.
5. A.A. Bolonkin, N.S. Khot, Optimal Structural Control Design, IAF-94-1.4.206, 45th Congress of the International Astronautical Federation, World Space Congress-1994. October 9-14,1994/Jerusalem, Israel.
6. V. Boltyanski, A. Poznyak, Linear multi-model time-optimization, Journal "Optimal Control Applications and Methods",Vol.23, Issue 3, 2002, pp.141-161.
7. Yunying Mao, Zeyi Liu, The optimal feedback control of the linear-quadratic control problem with a control inequality constraint, Journal "Optimal Control Applications and Methods",Vol.22, Issue 2, 2001, pp.95-109.
8. Heping Hua, Numerical solution of optimal control problems, Journal "Optimal Control Applications and Methods",Vol.21, Issue 5, 2000, pp.233-241.
9. H.Singh, R.H. Broun, D.S. Naidu, Unified approach to linear quadratic regulator with time-scale property, Journal "Optimal Control Applications and Methods",Vol.22, Issue 1, 2001, pp 1-16..
10. B.J. Driessen, N. Sadegh, Minimum-time control of systems with Coulomb friction: near global via mixed integer linear programming, Journal "Optimal Control Applications and Methods",Vol.22, Issue 2, 2001, pp.51-62.

Nomenclature

A is the $2n \times 2n$ plant matrix in liner problem

a_{ij} is members of matrix A

B is $2n \times p$ control matrix in liner problem

b_{jk} is members of matrix B

C is $q \times 2n$ out matrix

C_i are weight coefficients

c is constant

F_i is the magnitude of the bounds for each controller

F_0 is function of initial conditions

f is the control force vector of dimension p

$|f| \leq 1$ is the control in linear problem

H is Hamiltonian

I is the functional (objective function),

$\overline{M}, \overline{E}, \overline{K}$ are diagonal square matrices

P is a 2*n*-dimensional unknown matrix

$\underline{Q}$ is state weighting matrices

$\underline{R}$ is control weighting matrices

$R_x(t)$ is norm of displacement

T is final time

t is time (variable)

t_1, t_2, - boundary condition

u is the vector defines the structural response.

v is a *p*-dimensional vector of control forces

x is a *n*-dimensional vector of state in general problem,

x is the state vector of dimension *2n* in linear problem.

x(0) is the initial state vector

x(T) is the final state of the system.

x_1, x_2 - boundary condition

ζ is the vector of modal damping factors

$\lambda(t)$ is a *n*-dimensional vector unknown coefficient

λ is eigenvalues of matrix *A*

$\psi = \psi(t,x)$ is special function

ω is the vector of structural frequencies