CONJUGATE GRADIENT METHODS WITH AN APPLICATION
TO V/STOL FLIGHT-PATH OPTIMIZATION

By

R. K. Mehra and A. E. Bryson, Jr.

Technical Report No. 543

Reproduction in whole or in part is permitted for any purpose of the United States Government.

November 1967

The research reported in this document was made possible through support extended the Division of Engineering and Applied Physics, Harvard University by the U. S. Army Research Office, the U. S. Air Force Office of Scientific Research and the U. S. Office of Naval Research under the Joint Services Electronics Program by Contracts N00014-67-A-0298-0006, 0005, and 0008 and by the National Aeronautics and Space Administration under Grant NGR-22-007-068.

Division of Engineering and Applied Physics
Harvard University Cambridge, Massachusetts
CONJUGATE GRADIENT METHODS WITH AN APPLICATION TO V/STOL FLIGHT-PATH OPTIMIZATION

By

R. K. Mehra and A. E. Bryson, Jr.

Division of Engineering and Applied Physics
Harvard University Cambridge, Massachusetts

ABSTRACT

Conjugate gradient methods have recently been applied to some simple optimization problems and have been shown to converge faster than the methods of steepest descent. The present paper considers application of these methods to more complicated problems involving terminal as well as in-flight constraints. A number of methods are suggested to handle these constraints and the numerical difficulties associated with each method are discussed. The problem of flight-path optimization of a V/STOL aircraft has been considered and minimum time paths for the climb phase have been obtained using the conjugate gradient algorithm. In conclusion, some remarks are made about the relative efficiency of the different optimization schemes presently available for the solution of optimal control problems.
I. Introduction

Hestenes and Stiefel (1) in 1952 introduced the method of conjugate gradients for solving linear sets of equations. The same method was used by Fletcher and Reeves (2) in 1964 to solve nonlinear programming problems. Hayes (3) extended the method in 1954 to the solution of linear problems on Hilbert spaces. Antosiewicz and Rheinboldt (4) derived in 1962 convergence rates for these problems and showed that convergence is obtained in a finite number of steps for the linear-quadratic problem. Improved estimates of rates of convergence were obtained by Daniel (5) in 1965. Lasdon, Mitter and Warren (6) applied this method in 1966 to the solution of optimal control problems. They showed that the conjugate gradient method converged faster than the steepest-descent method on a number of problems. Sinnott and Luenberger (7) recently used another variant of the conjugate gradient method and gave similar results. In addition, they extended the method to handle linear terminal constraints.

Most of the optimal control problems solved so far (6, 7) using conjugate gradient methods have been simple in structure involving either no or very few terminal constraints. Lasdon, Warren and Rice (8) have tried using an extension of the Fiacco-McCormick "Sequential Unconstrained Minimization Technique" to handle in-flight inequality constraints, but the results were not too satisfactory for the problem of range-maximization of a re-entry vehicle. (This problem was originally solved by Bryson and Denham (9) using the method of steepest-descent). Speyer, Mehra and Bryson (10) solved
the same problem using a separation technique to handle the state-
variable inequality constraint. This separation method has been
described in detail in reference 10. It will be briefly outlined in
section IV along with other methods for handling in-flight constraints.
Some of these methods will then be applied in section V to the flight
path optimization of a V/STOL aircraft.

II. Conjugate Gradient Methods

a) Parameter Optimization: Conjugate gradient methods have the
property that they minimize a quadratic function of n variables in n
or less number of steps. They do so by generating a set of n directions
known as conjugate directions which span the n-dimensional space.
Let the function to be minimized be \( J = \frac{1}{2} (x - h)^T A (x - h) \) and let
\( p_0, p_1, \ldots, p_{n-1} \) be n vectors in Euclidean n space. They will be
called "A-orthogonal" or "A-conjugate", iff

\[
p_i^T A p_j = 0, \quad i \neq j
\]

where \( A \) is a positive definite matrix.

Therefore,

\[
p_i^T A p_i > 0, \quad \text{if } p_i \neq 0
\]

It is easy to show that n "A-conjugate" vectors are linearly inde-
dendent and form a basis for the n-dimensional space. If \( x_0 \) is the
initial guess, then \( (h - x_0) \) can be expressed in terms of this basis
as follows:

\[
h - x_0 = \sum_{i=0}^{n-1} a_i p_i
\]

where \( a_i = \frac{-p_i^T A_h}{p_i^T A p_i} = \frac{p_i^T g_0}{p_i^T A p_i} \)
where \( g_0 = A(x_0 - h) \) is the gradient vector \( \frac{\partial J}{\partial x} \bigg|_{x=x_0} \).

All conjugate gradient methods generate conjugate directions in one or another way. Basically, conjugate directions \( p_i \) can be generated by a Gram-Schmidt orthogonalization procedure starting from any arbitrary set \( v_0, v_1, \ldots, v_{n-1} \) of vectors. It can be shown that if \( v_i \) are the coordinate vectors, then the conjugate gradient method is functionally equivalent to the Gaussian elimination procedure. But the most convenient choice for \( r_i \) is the negative gradient vectors or the residue vectors \( r_i \):

\[
r_i = -g_i = A(h - x_i)
\]

This choice leads to a number of simplifications and, finally the following algorithm is obtained. Details of the proof can be found in Beckman (20).

\[
x_0 \text{ arbitrary}
\]

\[
g_0 = g(x_0)
\]

\[
P_0 = -g_0
\]

\[
x_{i+1} = x_i + a_i p_i \quad \text{where} \quad a_i = \frac{-p_i^T g_i}{p_i^T A p_i}
\]

\[
p_{i+1} = -g_{i+1} + \beta_i p_i \quad \text{where} \quad \beta_i = \frac{g_{i+1}^2}{g_i^2}
\]

This algorithm can be used for nonlinear programming problems as well. However, the matrix \( A \) is no longer a constant matrix and has to be computed at each step. One can avoid this by noting that if \( J \) is minimized along the direction \( (x_i + c_i p_i) \) with respect to \( c_i \), the optimum value of \( c_i \) is exactly \( a_i \) (20). Notice that if \( \beta_i = 0 \), the conjugate gradient method becomes a steepest descent method.
The conjugate gradient algorithm has a number of interesting properties. Rutishauser (11) compares it with other gradient methods and shows that it is the best method amongst a class of iterative gradient procedures for solving linear sets of equations. If \( \epsilon_i \) denotes the error vector \( \mathbf{h} - \mathbf{x}_i \), it can be shown that \( \| \epsilon_{i+1} \| < \| \epsilon_i \| \mathbf{V}_i \). Also, it can be shown that \( J \) is decreased at each step. Geometrically, \( p_i \) is the projection of the negative gradient vector \( g_i \) on to the subspace spanned by \( p_i, p_{i+1}, \ldots, p_{n-1} \). Thus we successively reduce the dimension of the subspace onto which \( -g_i \) is projected. This gives convergence in a finite number of steps.

b) Optimal Control Problems: Conjugate gradient methods can be readily extended to Hilbert spaces (3, 5). Consider the Mayer problem in the Calculus of Variations: Find \( u(t) \) to

minimize \( J = \phi(x(t_f)) \)

subject to \( \dot{x} = f(x, u, t) \) \hspace{1cm} (8)

\( x(t_0) \) and \( t_f \) are given, but \( x(t_f) \) is free.

\( x \) is an \( n \times 1 \) state vector and \( u \) is an \( r \times 1 \) control vector, both functions of time variable \( t \).

The Hamiltonian of the system is \( H = \lambda^T \mathbf{f} \), \hspace{1cm} (9)

and the adjoint equations are \( \dot{\lambda} = -f^\mathbf{T} \lambda \) \hspace{1cm} (10)

\[ \lambda(t_f) = \phi(\mathbf{x}(t_f)) \] \hspace{1cm} (11)

Let \[ g(t) = \frac{\partial H}{\partial u} = \lambda^T \frac{\partial f}{\partial u} \] \hspace{1cm} (12)
g is a vector of functions and relates $\delta J$ to $\delta u$ (13, chapter II).

$$\delta J = \int_{t_0}^{t_f} g \delta u \, dt$$

(13)

g plays the role of gradient vector in the finite dimensional case.

The same algorithm (equations 6 and 7) applies except that the scalar multiplications are changed to integrations. E.g., $\| g_i \|^2 = \int_{t_0}^{t_f} g_i^T g_i \, dt$.

c) Computation Details: A fourth order Runge-Kutta scheme is used to integrate the Euler-Lagrange Equations. It is necessary to store a direction of search to calculate the next direction of search. A cubic interpolation scheme (2) is used for one-dimensional search. It uses all the information available, i.e., $J(u_i), J(u_{i+1}), \frac{\partial J(u_i)}{\partial a_i}, \frac{\partial J(u_{i+1})}{\partial a_i}$ to fit the "smoothest curve" through the points $u_i$ and $u_{i+1}$, i.e., the curve which minimizes the integral $\int_0^{a_i} \frac{d^2 J}{d\sigma^2} \, d\sigma$ where $a_i$ is the step size.

III. Terminal Constraints

The conjugate gradient algorithm as given above applies only to unconstrained minimization problems. Modifications to the algorithm are necessary when there are constraints on the problem. A fairly general optimization problem with terminal constraints can be stated as follows: Find $u(t)$ to

minimize $J = \phi(x(t_f), t_f)$

subject to $\dot{x} = f(x, u, t) \ ; \ x(t_0)$ given
\[ \psi(x(t_f), t_f) = 0 \quad \text{q terminal constraints} \quad (16) \]

\[ \Omega(x(t_f), t_f) = 0 \quad \text{stopping condition for determining } t_f. \quad (17) \]

In effect, there are \((q + 1)\) terminal constraints. Any one of these can be chosen as a stopping condition. This is an unnecessary, arbitrary but useful device.

Two of the numerical methods for solving such problems are given below.

a) **Penalty Function Method:** Objective function \(J\) is modified using a quadratic penalty function

\[ \bar{J} = J + \psi^T K \psi \quad (18) \]

where \(K\) is a positive-definite matrix of penalty function constants. A sequence of unconstrained \(\bar{J}\) problems is solved with increasing values of \(K\). In the limit as \(K \rightarrow \infty\) we get \(\psi \rightarrow 0\), \(J \rightarrow J_{opt}\), \(u \rightarrow u_{opt}\). To check the efficiency of this method, it was used to solve a number of problems. The method worked quite well on linear-quadratic problems and simple nonlinear problems. Examples 1 and 2 of Ref. (6) were solved in one computer run by using a large enough value of \(K\). The minimum time earth-to-mars orbit transfer problem of Ref. (12) converged in 18 iterations starting from a stepped nominal and using about 1 minute of IBM 7094 computer time.

However, when this method was tried on flight path optimization problems involving aerodynamic drag and lift terms, the method ran into difficulties whenever the number of terminal constraints was increased beyond two. The reason seems to be that the "frozen-point" eigen-values of
the linearized system are split far apart due to damping terms. We define more clearly what we mean by eigen-values of a nonlinear system: If we linearize the equations around some nominal path and assume that the coefficients of the linearized equations vary sufficiently slowly in time that they may be considered constant over some period of time, we may talk about the eigen-values of this system. For a typical problem involving three state variables, \( V \) (velocity), \( h \) (altitude) and \( \gamma \) (flight path angle), the convergence was extremely slow if terminal constraints were put on all the three state variables simultaneously. Since for most of these problems, the terminal time is not specified, some sort of stopping condition is needed to determine \( t_f \) at each iteration. In this way, one of the constraints is automatically satisfied. It was found that the penalty function method could be used to handle at most two terminal constraints. If there were more terminal constraints, the convergence was extremely slow.

Various other types of penalty functions can be used. However, there is one common difficulty, viz., addition of penalty functions may change the problem completely creating narrow valleys and splitting the eigen-values of the system far apart. An example of this type of behavior is given in (13, chapter I). It is well-known that gradient procedures converge very slowly when the eigen-values of the system are split far apart.

b) Gradient Projection Method: Rosen's gradient projection method (14) was used by Bryson and Denham (9) to solve optimal control problems using a steepest descent method. The same method can be used with the conjugate gradient method to handle linear terminal
constraints (7). If the step size $a_i$ is small so that linearization is valid, the same method should work for nonlinear constraints as well. *

Bryson and Denham [9] have derived an expression for the projected gradient, $\bar{g}$. They show that

$$\bar{g} = f_u^T(\lambda_\phi - \lambda_\psi I_\psi I_\psi^T)$$

(19)

where

$$\lambda_\phi = -f_x^T \lambda_\phi$$

and

$$\lambda_\psi = -f_x^T \lambda_\psi$$

(20)

(21)

$$I_\psi^T = \int_{t_0}^{t_f} \lambda_\psi^T f_u^T f_u \lambda_\psi^T dt$$

(22)

$$I_\phi^T = \int_{t_0}^{t_f} \lambda_\phi^T f_u^T f_u \lambda_\phi^T dt$$

(23)

Conjugate directions $\bar{p}_i$ are generated using $\bar{g}_i$, $\bar{p}_{i-1}$ and equations (6) and (7). If a change $d\psi$ is desired in the constraint level, $\psi$, the control change $\delta u$ is given by

$$\delta u = f_u^T \lambda_\psi I_\psi I_\psi^T d\psi$$

(24)

The conjugate gradient algorithm is modified as follows:

$$\bar{u}_{i+1} = u_i + m_i a_i \bar{p}_i$$

(25)

(i) Start with $m_i = 1$ and obtain $a_i$ by a one-dimensional search.

(ii) Calculate the value of $\bar{\psi}_{i+1}(t_f)$ using $\bar{u}_{i+1}$. If linearization holds, $\bar{\psi}_{i+1}(t_f)$ should be the same as $\psi_i(t_f)$. If not, reduce $m_i$ so that $\|\bar{\psi}_{i+1} - \psi_i\| < \epsilon$, where $\epsilon$ is a small positive number.

* Authors do not have computational experience with this method so far.
(iii) Choose \( d\psi_1 \) and calculate the corresponding \( \delta u_1 \). Add this to \( u_{i+1} \)

\[
u_{i+1} = u_i + \delta u_1.
\]

Note that \( d\psi_1 \) should not be so large that the linearity assumption is violated.

If this algorithm is used on a linear-quadratic problem with linear terminal constraints, the directions of search \( p_i, i = 0, n - 1 \) will be conjugate and convergence would be obtained in a finite number of steps. For a nonlinear problem, however, the directions \( p_i, i = 0, n - 1 \), will not be conjugate in general due to the addition of \( \delta u \) from Eq. (24) at each step. To bypass this difficulty, one may try to satisfy the terminal constraints first and then hold them constant using the gradient projection scheme. This method would work well if the constraints are linear, but if the constraints are highly nonlinear, \( m_i \) will have to be chosen small enough so that linearization holds. In such a case, it might be better to approach near the optimum using the penalty function method and then refine the solution using the gradient projection method. Typically, in most of the optimization problems, the step size \( a_1 \) gets smaller and smaller as one approaches the optimum. So the linearization assumption would not be violated and the gradient projection method would generate conjugate directions near the minimum.

IV. In-Flight Constraints

There are three types of possible in-flight constraints which may be added to the problem statement in section III:

1. Control variable inequality constraints; \( N(u, t) \leq 0 \).  
   \hspace{2cm} (26)

2. Control and state variable inequality constraints or mixed constraints; \( C(x, u, t) \leq 0 \).  
   \hspace{2cm} (27)

3. State variable inequality constraints; \( S(x, t) \leq 0 \).  
   \hspace{2cm} (28)
We first describe general methods (a), (b) and (c) applicable
to all of these cases. Then we describe special methods (d) and (e) for
particular types of constraints.

a) **Penalty Function Methods**

Consider a scalar mixed constraint

\[ c(x, u, t) \leq 0. \]  \hfill (29)

Introduce a new state variable \( r \) such that

\[ \begin{cases} 
   Kc^2 & \text{if } c > 0 \\
   0 & \text{if } c \leq 0 
\end{cases} \]  \hfill (30)

where \( K \) is a large positive constant and

\[ r(t_0) = 0. \]

Then if \( r(t_f) \geq 0 \), the constraint \( c \leq 0 \) is approximately satisfied.

The Interior Penalty Method of (8) tries to solve a sequence of
minimization problems

\[ P[u, r] = \phi(x(t_f), t_f) - r \int_{t_0}^{t_f} \frac{dt}{c(x, u, t)} \]  \hfill (31)

where \( r \) is a positive scalar and tends to zero. It can be shown that
this method approaches the constraint boundaries from the interior (8).

Our experience has shown that these methods work poorly on
highly nonlinear control problems. Lasdon et al. (8) encounter con-
siderable difficulty in solving the re-entry problem. Moreover the
constraint can never be exactly satisfied because these methods work
by violation of the constraints.
b) Transformation of Variables

M. J. Box (15) has used this method for solving some nonlinear programming problems. It can be used for optimal control problems also, e.g., if a variable $S$ has to be positive, we can use another variable $y$ which is unconstrained and is related to $S$ as

(i) $S = y^2 \quad (32)$

or

(ii) $S = e^y \quad (33)$

Similarly, if $0 \leq S \leq 1$

(i) $S = \sin^2 y \quad (34)$

or

(ii) $S = \frac{e^y}{e^y + e^{-y}} \quad (35)$

If $S_{\text{min}} \leq S \leq S_{\text{max}}$,

$$S = S_{\text{min}} + (S_{\text{max}} - S_{\text{min}})\sin^2 y \quad (36)$$

Then an unconstrained problem in $y$ space is solved.

These methods are applicable to control and mixed type constraints only. Moreover, they produce slow convergence near the boundaries, e.g., if $S_{\text{min}} \leq S \leq S_{\text{max}}$ and the above transformation is used,

$$\frac{\partial S}{\partial y} = (S_{\text{max}} - S_{\text{min}})\sin 2y \quad (37)$$

which will be zero for

$$y = 0, \frac{\pi}{2} \quad \text{or for } S = S_{\text{min}}, S_{\text{max}}.$$

c) Gradient Projection Method

Bryson, Denham, and Dreyfus (16) have shown that inequality constraints can be handled by solving for a set of inner-point equality constraints plus control constraints. Gradient projection can be used
with the conjugate gradient method to handle inner point constraints in the same way as the terminal constraints. The gradient vector \(g\) is computed using equations given by Bryson and Denham (9). The conjugate directions are generated separately for the paths before and after the inner-point constraint. The gradient on the arc before the inner-point constraint is projected on the intersection of two subspaces viz. those of the inner-point constraints and the terminal constraints whereas the gradient on the arc after the inner-point constraint is projected on one subspace only viz. subspace of the terminal constraints.

d) **Control Variable Inequality Constraints: Bang-Bang Solution**

If the control variable enters linearly in the equations of motion and in the performance index of the problem, one can show (13) that the control always lies on one of the boundaries. The problem is thus reduced to determining the switching times.

If \(u_{\min} \leq u \leq u_{\max}\) and if \(t_i\)'s are the switching times, then

\[
\frac{\partial J}{\partial t_i} = (-1)^{i-1}(\lambda^T f_u)(u_{\max} - u_{\min})
\]  

(38)

assuming at \(t_1\), \(u\) goes from \(u_{\max}\) to \(u_{\min}\) (9).

Treating \(t_i\) as control parameters, we can iterate on them to obtain optimum switching times. If the number of switching times is unknown, it is better to start with more switching times than anticipated. The above technique can eliminate some switching times, but it cannot add extra switching times.

Then parameter optimization can be carried out using the conjugate gradient method (section IIa).
e) **State Variable Inequality Constraints (SVIC); Separation of Arcs**

This method is due to Speyer (17). It is applicable whenever the order of the SVIC is one less than the number of state variables in the system. In such cases, it becomes possible to compute the unconstrained arcs separately (10). One finds that the motion along the constraint boundary depends only on one state variable. Using this as the variable of integration, the value of the performance index along the constraint boundary is expressed as a function of the entry and exit point values of this variable. These functions are lumped suitably with objective functions along unconstrained arcs and the problem is reduced to a set of unconstrained problems. Speyer, Mehra and Bryson (10) use this method to solve the problem of range-maximization of a re-entry glider (9). The results obtained by them show that this method is very powerful, whenever it can be applied to problems with SVIC.

**V. Flight Path For Minimum Time Climb-To-Cruise of a V/STOL Aircraft**

Compared to conventional aircraft, V/STOL aircraft have an extra control variable, namely the angle between the thrust direction and a reference axis in the aircraft. It is of interest to know how this extra control variable may be used to improve the performance of the aircraft.

If a flight is long enough, it can be divided into three paths:

(i) Climb phase starting from the ground and going up to some cruise condition;

(ii) Cruise at some constant altitude and velocity;

(iii) Landing phase.
Depending on the particular use to which the V/STOL aircraft is put, there may be flight path constraints on (i) and (iii).

If the cruise conditions are known, the optimization problem reduces to optimization of the two arcs (i) and (iii) separately, because the cruise conditions specify the state completely at the end of path (i) and at the beginning of path (iii).

Here, we shall consider the hypothetical jet-lift aircraft to Ref. (18).* Gallant (19) has considered a tilt-wing V/STOL aircraft and obtained minimum-direct-cost flight paths for a 50 mile flight starting from the end of the transition to the beginning of the retransition.

Problem Formulation

The aircraft will be approximated as a mass-point. Figure 1 shows the forces acting on the aircraft. Figure 2 shows the thrust force in greater detail. It is assumed that the jet inlets are always pointed in the direction of the relative wind velocity. This approximation is reasonable in view of the rough model assumed for the V/STOL aircraft and in view of the final results which show that the angle-of-attack is kept small during most of the flight.

The equations of motion are:

\[ \dot{V} = \frac{T}{m} \cos(a+i) - \frac{D}{m} - g \sin \gamma - \frac{M}{m} V[1 - \cos(a+i)] \]  (39)

\[ \dot{\gamma} = \frac{T}{mV} \sin(a+i) + \frac{L}{mV} - \frac{g}{V} \cos \gamma + \frac{M}{m} \sin(a+i) \]  (40)

\[ \dot{h} = V \sin \gamma \]  (41)

\[ \dot{x} = V \cos \gamma \]  (42)

* Authors gratefully acknowledge the help and suggestions received from Professor R. H. Miller and his students at the Flight Transportation Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts.
\[ L + D_m' \]

\[ D_m = \text{Momentum Drag Along V-axis.} \]

\[ D_m' = \text{Momentum Drag \perpendicular \text{To V-axis.}} \]

**FIGURE 1**

\[ \text{FIGURE 2} \]

\[ \dot{M}, V_e \]

\[ \text{RATE OF AIR FLOW} = \dot{M} \]

\[ \text{EXIT VELOCITY OF GASES} = V_e \]

Force along V axis = \[ F_V = \dot{M}V_e \cos (\alpha+i) - \dot{M}V \]

Force \perp \text{to V direction} = \[ F_Y = \dot{M}V_e \sin (\alpha+i) \]

**Thrust** \[ T = \dot{M}V_e - \dot{M}V \text{ (Equal net force when (\alpha+i) = 0)} \]

\[ \therefore F_V = T \cos (\alpha+i) - \dot{M}V (1 - \cos(\alpha+i)) \]

\[ F_Y = T \sin (\alpha+i) + \dot{M}V \sin (\alpha+i) \]
Where

\[
L = \frac{1}{2} \rho V^2 C_L S \quad (43)
\]

\[
D = \frac{1}{2} \rho V^2 C_D S \quad (44)
\]

\[
C_L = C_{L_a} a \quad (45)
\]

\[
C_D = C_{D_o} + C_{D_1} a^2 \quad (46)
\]

Air Density \( \rho = 0.0023769(1 - 0.6875 \times 10^{-5} h)^{4.2561} \quad (47) \)

Equation (47) holds for \( h \leq 36,000 \) ft.

The characteristics of the hypothetical aircraft are

Thrust \( T = T_o \left( 1 - \frac{55h}{30,000} \right) \) where \( h \) is in ft.

Mass \( m = \frac{56902}{32.2} \) slugs (taken as constant during climb)

Wing Area \( S = 421 \) ft\(^2\)

\( C_{D_o} = .027 \)

\( C_{L_a} = 5.73 \)

\[
C_{D_1} = \frac{C_{L_a}^2}{\pi e AR} = \frac{(5.73)^2}{\pi \times 0.9 \times 6} = 1.93
\]

Rate of Air Flow \( \dot{M} = \frac{T_o}{(65 \times 32.2)} \) slugs/sec.

if \( T_o \) is in lb.

There are three control variables in the problem:

Magnitude of thrust vector \( (T_o) \), \( 0 \leq T_o \leq T_{o \text{ max}} \);

Direction of thrust vector \( (i) \);

Angle of attack \( (a) \) or pitch angle \( (\theta) \).
It is preferable to use $\theta$ instead of $\alpha$ as the control variable. The use of $\theta$ as the control variable adds extra damping terms into the equations of motion which help in convergence.

We shall obtain minimum time paths under the following assumptions:

1) **Thrust**, $T_0$, is kept constant at its maximum value. This is a reasonable assumption for the climb phase of the flight. In particular, we shall use $T_0 = 1.25 mg$.

2) **Initial conditions** for the problem are

$$V(0) = 0, \quad h(0) = 0, \quad x(0) = 0$$

The $\dot{V}$ equation has a singularity at $V = 0$. To integrate the equations of motion numerically, we must start with a finite $V$. The aircraft would attain this velocity after flying for some time, say $t_1$, in some particular manner. This part of the flight may be partially or completely determined by restrictions on the runway available for take-off, e.g., if the aircraft must take-off vertically, then $\gamma(t_1) = 90^\circ$ where $t_1$ will be some time either during or at the end of the vertical take-off period. We will now consider a few specific cases.

**Unconstrained Take-Off.** To get an idea as to what the aircraft should do if there were no constraints imposed on it due to the ground, we shall consider the case in which the aircraft can even go underground. While this case is unrealistic, it will provide useful information about the optimal paths with constraints. We do not know whether a V/STOL aircraft should take-off like a conventional aircraft (by first picking up speed along the runway) or whether it should take-off directly making some angle $\gamma(0+) > 0$.
to the horizontal. If we solve the unconstrained problem treating \( Y(t_1) \) as a control parameter, we should be able to answer this question.

The initial conditions for this case may, therefore, be taken as, (treating \( t_1 \) as starting time denoted by 0)

\[
V(0) = 50 \text{ ft/sec} \\
Y(0) \text{ chosen to make } \lambda_Y(0) = 0 \\
\]

(This means that the optimization process must drive \( \lambda_Y(0) \) to zero.)

\[
h(0) = 0 \quad x(0) = 0 \\
\]

Changing \( h(0) \) from 0 to several hundred feet will not change the results significantly.

The terminal conditions are the cruise conditions. The final time, \( t_f \), is to be minimized:

\[
\gamma(t_f) = 0 \\
h(t_f) = 20,000 \text{ ft} \\
V(t_f) \text{ free} \\
x(t_f) \text{ free} \\
\]

A constraint on \( V(t_f) \) could be met easily either by changing the path slightly or by changing thrust magnitude towards the end of the climb phase. The control variables used are \( \theta \) and \( i \).

Figures 3, 4, 5a, 6, 7, and 8 show the results obtained for the case when there are no constraints on take-off. The optimum value of \( \gamma(0) \) at \( V = 50 \text{ ft/sec} \) turns out to be about 7°. But the interesting fact is that \( \gamma \) soon becomes negative and the aircraft goes about 300 ft underground. Reasons for this seem to be:
FIG. 4 VELOCITY vs. RANGE FOR (i) UNCONSTRAINED TAKE-OFF (ii) VERTICAL TAKE-OFF
FIG. 5a  FLIGHT-PATH ANGLE, $\gamma$, $(\alpha+i)$, AND THRUST DIRECTION, $\alpha+i+\gamma$, VS. TIME FOR UNCONSTRAINED TAKE-OFF.
FIG. 5b FLIGHT-PATH ANGLE AND (a+i) VS. TIME WITH VERTICAL TAKE-OFF CONSTRAINT.
FIG. 6 PITCH ANGLE, $\theta$, VS. TIME FOR (i) UNCONSTRAINED TAKE-OFF (ii) VERTICAL TAKE-OFF.
FIG. 7 JET-TILT ANGLE, \( i \), vs. TIME FOR (i) UNCONSTRAINED TAKE-OFF (ii) VERTICAL TAKE-OFF.
FIG. 8 ANGLE-OF-ATTACK VS. TIME FOR (i) UNCONSTRAINED TAKE-OFF (ii) VERTICAL TAKE-OFF
(i) The thrust is greater at lower altitudes;

(ii) The aircraft should pick up velocity as fast as possible in order to generate aerodynamic lift. Lift obtained from tilting the jet is not very efficient because it gives lots of momentum drag. The angle (a+i) is apparently kept low in order to keep this drag low (cf. Fig. 5a).

To keep the surface drag low, angle of attack \( \alpha \) is also kept small as shown in Fig. 8. The aircraft dives down because gravity helps it in picking up speed. Figure 4 shows velocity vs. range.

After the aircraft has picked up velocity during the diving maneuver, \( \gamma \) increases quickly to a maximum value of 56.6\(^{\circ}\) (Fig. 5a). Figure 3 shows how \( h \) increases during this phase. However, \( \theta \) also increases at the same time so that \( \alpha = \theta - \gamma \) remains small. Jet-tilt angle \( i \) is also kept small. Thus, the total drag is kept low. This maneuver is followed by a rapid change in \( \theta \) to a negative value of about -25\(^{\circ}\). This is necessary to meet the terminal condition on \( \gamma \) viz. \( \gamma(t_f) = 0 \). The total time taken by the aircraft is 53 sec. Calculations show that if the aircraft is made to climb vertically all the way up from the ground, it takes twice as much time. The velocity in that case never exceeds 300 ft/sec.

Thus the results show that a V/STOL aircraft without take-off constraints should fly very much like a conventional aircraft. Aerodynamic lift is more efficient than jet-lift. On the other hand, the aircraft should keep angle-of-attack \( \alpha \) small to keep aerodynamic drag low.
Horizontal Take-Off Constraint. Let us now impose the constraint that the aircraft cannot go underground. However, it can still run along the ground. Numerical results show that when the aircraft reaches $V \approx 300 \text{ ft/sec}$, the optimal path no longer dives down. We could integrate the equations of motion of the aircraft on the ground to calculate the distance in which the aircraft can attain this velocity. This also indicates that the best take-off for a STOL aircraft is to fly parallel to the ground at low altitude for a considerable distance.

Vertical Take-Off Constraint. We now impose the restriction that the aircraft must fly vertically up to an altitude of 1000 feet. From the results obtained above, it appears that the best way to do this would be to make $\theta = 90^\circ$ so that $a = 0$, $i = 0$, $\gamma = 0$. Integration of the $V$ equation gives $V = 125 \text{ ft/sec}$ at $h = 1000 \text{ ft}$. Time taken is 8 seconds. The optimization problem is now solved with the following initial conditions:

\[
V(8) = 125 \text{ ft/sec} \\
\gamma(8) = 90^\circ \\
h(8) = 1000 \text{ ft.} \\
x(8) = 0
\]

The results are shown in Figs. 3, 4, 5b, 6, 7, and 8. The total path (from take-off) is 60 seconds long and is similar to the unconstrained take-off case. The aircraft goes up first due to positive $\gamma$, but soon dives down to a minimum altitude of about 980 ft. The control variables $\theta$ and $i$ have discontinuities at $t = 8 \text{ sec}$ when the constraints are relaxed.

Similar behavior would be obtained if the aircraft were constrained to take-off at some other constant value of flight path angle.
\(\gamma_C\). The equation \(\dot{\gamma} = 0\) determines one of the control variables in terms of the other (say \(i\) in terms of \(\theta\)). If \(\theta\) is constrained by \(\theta \leq \gamma_C\), as in the case above, one would intuitively expect that \(\theta\) would remain constant at \(\gamma_C\).

VI. Conclusion

Our computational experience has shown that the conjugate gradient method, though very efficient for simple optimization problems, may run into difficulties when applied to more complicated problems. Some of the difficulties that may be encountered are:

1. Gradient of the objective function with respect to the step size may not become zero or small enough during one-dimensional search. Accumulation of errors due to this source can produce directions of search which increase rather than decrease the performance index. In such cases, it was found useful to revert back to the local gradient direction and start the process over again. This procedure is similar to the one suggested by Beckman (20) and also used by Fletcher and Reeves (2) for nonlinear programming problems.

2. Use of penalty functions may split the "frozen-point" eigen-values of the linearized system far apart and make convergence extremely slow. The use of the gradient projection method, though more complicated, may help in this case, particularly near the optimum.

Conjugate Gradient Methods vs. Steepest Descent Methods:

(i) For optimal control problems having either no or few constraints, conjugate gradient methods, though requiring more programming, are faster and lead to a better solution than steepest descent methods. For linear-quadratic problems, conjugate gradient methods reach the optimum in a finite number of steps.
(ii) For optimal control problems with a large number of constraints, conjugate gradient methods run into the above-mentioned difficulties. It then becomes necessary to control the step size during the one-dimensional search and the directions of search are no longer conjugate to each other. In nonlinear problems with nonlinear constraints, it is not clear that conjugate gradient methods would do better than steepest descent methods except when starting close to the optimum.

Conjugate Gradient Methods vs. Second Variation Methods:

(i) Conjugate gradient methods require less programming and less computation per iteration than second variation methods.

(ii) Second variation methods require the matrix of second variations of the Hamiltonian with respect to the control \( \mathbf{H}_{uu} \) to be non-singular. Conjugate gradient methods do not require this.

(iii) Conjugate gradient methods do not converge to extremals containing conjugate points, whereas second variation methods try to converge towards these extremals.

(iv) Second variation methods lead to more accurate solutions than conjugate gradient methods, particularly to better control histories.
References


Conjugate gradient methods have recently been applied to some simple optimization problems and have been shown to converge faster than the methods of steepest descent. The present paper considers application of these methods to more complicated problems involving terminal as well as in-flight constraints. A number of methods are suggested to handle these constraints and the numerical difficulties associated with each method are discussed. The problem of flight-path optimization of a V/STOL aircraft has been considered and minimum time paths for the climb phase have been obtained using the conjugate gradient algorithm. In conclusion, some remarks are made about the relative efficiency of the different optimization schemes presently available for the solution of optimal control problems.
<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>LINK A</th>
<th>LINK B</th>
<th>LINK C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimization</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conjugate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gradient</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>