OPTIMUM CONFIGURATIONS FOR BANGLESS SONIC BOOMS*

BY

WALLACE D. HAYES AND FRANCIS B. WEISKOPF, JR.

Princeton University

Abstract. A number of optimization problems are posed and solved for supersonic aircraft flight subject to the condition that a shock wave appears only incipiently in the sonic boom signal at a given point. The principal result is one giving the maximum effective gross weight of an aircraft of given effective length under given flight conditions. The calculus of variations with inequality constraints is used, with the novel features of a non-local isoperimetric relation and of only an upper bound on a control variable.

1. Introduction. Both the term "sonic bang," used in Great Britain, and the term "sonic boom", used in the United States, refer to the same phenomenon, the quasilinear acoustic signal emitted by an aircraft moving supersonically. To resolve partly this dichotomy of terminology and to make the terms quantitative, it is proposed that the bang should be the pressure jump across the strongest shock wave in the signal and that the boom should be the impulse $\int p \, dt$ in the primary identifiable lobe of the signal [1]. The bang then is a measure of the major element of psychoacoustic annoyance; the boom is a measure (more or less) of the potential for structural damage, and is in a certain sense independent of nonlinear distortion of the signal.

One potential method of attaining commercial supersonic flight overland on a basis acceptable to voting inhabitants underneath the flight paths is to use aircraft which produce no sonic bang on the ground in normal operations. This method was suggested by McLean [2]. Such aircraft would produce a bangless sonic boom, one with no shocks in the signature. This can be accomplished by designing the aircraft with appropriate distributions of lifting and volume elements. How the aircraft is to be so designed is a problem in supersonic aerodynamics and is not the subject of this paper. Here we consider the aircraft to be represented by a fore-and-aft line distribution of either lift or volume elements, and refer to this distribution as the configuration. We assume linear theory to apply near the aircraft; where nonlinear effects are locally important, the treatment of these effects is considered part of the supersonic aerodynamic design problem.

The purpose of this paper is to establish optimum configurations subject to the bangless boom condition and various other isoperimetric conditions. The effective length of the aircraft is always considered fixed and, in most problems considered, the lift on the aircraft is to be maximized.

The problems considered are ones in the calculus of variations with inequality constraints. The solutions are divinable on heuristic grounds by one familiar with this

^{*} Received August 27, 1971. The support of N.A.S.A. under Grant NGL 31-001-119 is gratefully acknowledged.

particular subject. The now-standard methods using switching functions require some extension to handle the present problems, but they do apply and thereby establish the results without guesswork. These solutions, as obtained through the calculus of variations, are the principal subject of this paper.

The problem posed in the calculus of variations is unconventional in two ways: one of the isoperimetric relations connecting two dependent variables is an Abel integral transform and is thereby a nonlocal relation; only an upper bound is imposed upon the control variable (F'), whose integral may have negative discontinuities. One feature of the problem is that the specific solution for the switching function is not needed in order to solve for the variables of physical interest. The most useful information obtained from the switching function is the maximum number of zeros it may have.

2. Basic sonic boom relations. We here consider only the simplest realistic case, that of the sonic boom from an aircraft in uniform level flight in a stratified atmosphere. For the most part we consider only the case in which the atmosphere is isothermal, and only the signal immediately downward from the aircraft (at azimuth angle $\phi = 0$). This simplification does not restrict the optimization analysis in any way, and is intended to yield an adequate basis for the optimization and to provide useful formulas. These formulas can be modified to fit, for example, propagation at other azimuth angles in other than isothermal atmospheres, with winds if desired.

The approach we shall use is that of [3]. In this approach a suitable phase variable and invariant signal strength variable are identified, and an age variable appropriate to these identified quantities is defined for the purpose of calculating the distortion of the signal. With the simplifications made we may use the natural argument $x - \beta r$ of the Whitham F-function as phase variable, relabelled x as though all calculations were made on the aircraft axis. This variable represents distance aft of a reference Mach cone, measured on a line parallel to the aircraft flight axis. The quantity x divided by the aircraft speed U is the time phase ξ of [3]. For the invariant signal-strength variable it is possible in this case to use the F-function itself, obtained from the aircraft shape and aerodynamic loads by the method of oblique planes. For a discussion of the literature and a review of the concepts involved, see [1] or [3].

The azimuth angle ϕ is the angle variable in a cylindrical coordinate system aligned with the aircraft flight axis, with the value $\phi = 0$ corresponding to the direction straight down to the ground. The aircraft is represented for each value of ϕ either by a line lift distribution $f(x, \phi)$ per unit distance or by a line source distribution described by the cross-sectional area $S(x, \phi)$ of an equivalent slender body of revolution. The two distributions are related by

$$f = \rho_a U^2 \beta^{-1} S', \tag{2.1}$$

where the prime indicates differentiation with respect to x and $\beta = (M^2 - 1)^{1/2}$. Here $M = U/a_a$, and ρ_a and a_a are the atmospheric density and sound speed near the aircraft. The Whitham F-function is defined

$$F(x,\phi) = \frac{1}{2\pi} \int_0^x \frac{S''(\xi,\phi)}{(x-\xi)^{1/2}} d\xi = \frac{\beta}{2\pi\rho_a U^2} \int_0^x \frac{f'(\xi,\phi)}{(x-\xi)^{1/2}} d\xi$$
 (2.2)

It has been defined so that the pressure perturbation p and velocity perturbation q in a uniform atmosphere far from the aircraft are given by

$$(1/M^2)(p/\rho a^2) = (1/M^2)(q/a) = (1/2\beta r)^{1/2}F, \qquad (2.3)$$

according to linear aerodynamic or acoustic theory. Here r is the radial cylindrical coordinate, distance from the aircraft flight axis.

In an atmosphere with variable properties without winds the invariant according to linear geometric acoustics is $\rho aq^2 \alpha = \rho^{-1}a^{-1}(\Delta p)^2 \alpha$, where α is a suitable measure of ray tube area. With the simplifications of uniform level flight in an isothermal atmosphere we may take $\alpha = r$. If we now replace (2.3) by

$$F = ((2\beta)^{1/2}/M^2)(\rho a^3/\rho_a a_a^3)^{1/2}(q/a)r^{1/2}, \tag{2.4}$$

then F is itself the desired invariant variable. In the isothermal atmosphere the factor $(a/a_a)^{3/2}$ disappears.

The nonlinear distortion appears because the signal travels faster than the speed of sound by the quantity $\frac{1}{2}(\gamma + 1)q$. With the distance s normal to the reference Mach cone or wave front used as a phase, this effect gives a phase shift

$$ds/dt = -\frac{1}{2}(\gamma + 1)q. {(2.5)}$$

But x = Ms from the geometry of the Mach cone, while $dr = \beta M^{-1}a \ dt$. Substituting from these relations and from (2.4), (2.5) may be put in the form

$$dx/d\tau = -F, (2.6)$$

where

$$\tau = \frac{\frac{1}{2}(\gamma + 1)M^4}{2^{1/2}\beta^{3/2}} \int_0^4 \left(\frac{\rho_a}{\rho}\right)^{1/2} r^{-1/2} dr \tag{2.7}$$

is termed the age. The integral is an indefinite one evaluated to be zero at the indicated lower limit.

The integral of Eq. (2.6) is $\delta x = -\tau F$, and represents the phase shift of a part of the signal of strength F. The inverse slope of the F-function is changed by the distortion from $dx/dF = (F')^{-1}$ to $(F')^{-1} - \tau$. A shock first appears in the signal where this quantity first becomes zero. Our basic requirement is that a shock does not appear in the signal before the signal hits the ground. With τ interpreted as the value at the ground this condition is

$$\sup F' = A \le \tau^{-1}. \tag{2.8}$$

The condition that no shock appear at the ground is the inequality of (2.8). For the purposes of the analysis we assume that a shock just appears at the ground, with the equality of (2.8) satisfied. This condition could be easily replaced by a condition, for example, on the maximum value of dp/dt permitted in the signal at the ground.

A property of an isothermal atmosphere (of uniform composition) is that the density obeys an exponential law. In the present case, the density is given by

$$\rho/\rho_a = e^{\alpha r}, \tag{2.9}$$

where the inverse scale distance is

$$\alpha = \gamma g \cos \phi / a_a^2 . \tag{2.10}$$

With this law the age τ may be expressed from Eq. (2.7) as

$$\tau = \frac{\frac{1}{2}(\gamma + 1)\pi^{1/2}M^4 \operatorname{erf}(\zeta)}{\alpha^{1/2}\beta^{3/2}} = \frac{(\gamma + 1)\pi^{1/2}M^4a_0 \operatorname{erf}(\zeta)}{2\gamma^{1/2}\beta^{3/2}q^{1/2}(\cos\phi)^{1/2}},$$
 (2.11)

where

$$\zeta = (\frac{1}{2}\alpha r)^{1/2} = \left[\frac{1}{2}\gamma g a_a^{-2} (z_a - z)\right]^{1/2} \tag{2.12}$$

and $z_a - z$ is the altitude of the aircraft above the ground. The function erf (ζ) is

erf
$$(\zeta) = \frac{2}{\pi^{1/2}} \int_0^{\zeta} \exp(-\zeta'^2) d\zeta'$$
.

The aircraft is characterized by an effective length L and an effective lift or weight W. The effective length of the aircraft is the range of x starting at x = 0 over which f can be non-zero. The effective weight of the aircraft is the actual weight plus $\rho_a U^2 \beta^{-1}$ times the increase in exit area over capture area for the aircraft engines. These entities are discussed further in [1]. For our purposes we have the relation

$$W = \int_0^L f(x) \ dx \tag{2.13}$$

as basic to our problem. We will in certain problems deal with the volume V of a slender body of revolution, given by

$$V = \int_0^L S(x) dx = \int_0^L (L - x)S'(x) dx, \qquad (2.14)$$

with the condition S(0) = 0 assumed. The moment M of a loading and the location x_c of the center of the loading are defined by

$$M = Wx_{\epsilon} = \int_0^L x f(x) dx. \qquad (2.15)$$

With the equivalence between lift and source distributions the identity $M + \rho_a U^2 \beta^{-1} V = LW$ holds.

As a step preliminary to the optimization analysis the quantities A, W, V, M, and x_c and the variables x, F(x), f(x), and S(x) are reduced to dimensionless form by

$$A = L^{-1/2}A^*, W = \rho_a U^2 \beta^{-1} L^2 W^*, V = L^3 V^*,$$

$$M = \rho_a U^2 \beta^{-1} L^3 M^*, x_e = L x_e^*, x = L x^*,$$

$$F = L^{1/2} F^*(x^*), f = \rho_a U^2 \beta^{-1} L f^*(x^*), S = L^2 S^*(x^*).$$

$$(2.16)$$

The resulting equations for dimensionless quantities are the same as the corresponding dimensional ones with L=1 and $\rho_a U^2 \beta^{-1}=1$. In the optimization analysis these reduced quantities are used, but with the asterisks dropped. We return to dimensional quantities in the final section, where the results are discussed.

3. Basic optimization theory. In terms of the dimensionless quantities and variables the range of x for which f(x) = S'(x) can be nonzero is $0 \le x \le 1$. In all cases we require that $F' \le A$. We may wish to maximize $W = \int_0^1 f \, dx$, or we may wish to maximize $V = \int_0^1 (1-x)f \, dx$. We may wish to impose isoperimetric conditions on M with W maximized (to fix center of pressure for a lift loading). We may wish to impose an isoperimetric condition on S(1) = W in the case in which V is maximized.

We uniformly impose the boundary conditions at x = 0

$$S(0) = 0, (3.1)$$

$$f(0) = S'(0) = 0. (3.2)$$

Eq. (2.2) now has the form

$$F(x) = \frac{1}{2\pi} \int_0^x f'(\xi) \, d\xi / (x - \xi)^{1/2}. \tag{3.3}$$

This is an Abel integral equation whose solution with boundary conditions (3.1) and (3.2) is

$$f(x) = 2 \int_0^x F(\xi) d\xi / (x - \xi)^{1/2}. \tag{3.4}$$

Eqs. (3.3) and (3.4) are to be treated as equivalent.

The function F must satisfy the basic inequality $F' \leq A$ in the range x > 1 also. A discontinuity in f(x) would give an unacceptable $(-\frac{1}{2})$ -power singularity in F(x). Since f(x) = 0 for x > 1 we impose uniformly the boundary condition

$$f(1-) = 0. (3.5)$$

We treat formally two classes of problems, that of our basic *single-bound* case with one bound on F' and a second one of a *double-bound* case. In either case we must impose additional constraints analogous to that of (3.5), arising from the requirement that $F'(1 + \epsilon)$ satisfy for small ϵ the same condition that F'(x) does in the interval 0 < x < 1.

In the first class of problems we impose:

single-bound case:
$$F'(x) \le A$$
, (3.6)

$$F(0+) = -a^2 \le 0, (3.7)$$

$$f'(1-) = b^2 \ge 0, (3.8)$$

$$f''(1-) \ge 0$$
 if $b^2 = 0$, $F'(1-) = A$. (3.9)

In the second class of problems we also impose a lower limit -D on F', and impose

double-bound case:
$$-D \le F'(x) \le A$$
, (3.10)

$$F(0+) = 0, (3.11)$$

$$f'(1-) = 0, (3.12)$$

$$f''(1-) \ge 0$$
 if $F'(1-) = A$, (3.13a)

$$f''(1-) \le 0$$
 if $F'(1-) = -D$. (3.13b)

The double-bound case should approach the single-bound case in the limit $D \to \infty$, and in this property lies our reason for considering the second case. The double-bound case could be of some physical interest in its own right, but we choose not to explore this possibility.

The problem in the calculus of variations is established using F(x) and f(x) as dependent variables. Either (3.3) or (3.4) must be posed as an isoperimetric constraint; we choose to use (3.4). Conditions (3.5) and (3.8) or (3.12) are used both as boundary conditions on f and also as isoperimetric constraints on F. Conditions (3.9) or (3.13) are not put directly into the problem but are checked after a solution has been obtained.

The basic inequality constraint (3.6) or (3.10) is treated by the standard approach ([4] or [5]) of the calculus of variations with inequality constraints. A variable z (sometimes termed the "slack" variable) is introduced, defined for constraint (3.10) by

$$z^{2} = (A - F')(1 + (F'/D)) \ge 0 \tag{3.14}$$

wherever F' is defined. The variable z is constrained to be real. Constraint (3.6) is obtained by setting F'/D = 0 in (3.14). The role of z is to permit F' to be other than A or -D in the solution. In fact, it turns out that F' never takes other than these values in the range 0 < x < 1, and z turns out to be zero except at discrete points.

The calculus of variation problem posed is that of minimizing $I = \int_0^1 G \, dx$, where G includes the negative of the quantity to be maximized, plus Lagrange constant multipliers times quantities whose integrals are to be held constant, plus Lagrange multipliers times relations to be set to zero over the range 0 < x < 1.

The integral of G includes the integral $\int_0^1 dx$ of a variable multiplier $\lambda_4(x)$ times the expression $[2 \int_0^x (x - \xi)^{-1/2} F(\xi) d\xi - f(x)]$, which from Eq. (3.4) must equal zero. This equation is a nonlocal relation between f and F. This unconventional feature is handled by the replacement

$$2\lambda_4(x) \int_0^x \frac{F(\xi) \ d\xi}{(x-\xi)^{1/2}} \to 2F(x) \int_x^1 \frac{\lambda_4(\xi) \ d\xi}{(\xi-x)^{1/2}}$$
 (3.15)

obtained by interchanging the order of the double integration.

With this replacement we write G as

$$G(x) = \lambda_0 f(x)(1-x) + \lambda_1 f(x) + \lambda_2 \frac{2F}{(1-x)^{-1/2}} + \lambda_3 \left[\frac{2F'}{(1-x)^{1/2}} - b^2 \right] + 2F \int_x^1 \frac{\lambda_4(\xi) d\xi}{(\xi-x)^{1/2}} - \lambda_4 f + \lambda_5(x)[(A-F')(1+F'/D) - z^2].$$
(3.16)

This formulation is sufficiently general to encompass the cases to be considered. The terms in the constants λ_0 and λ_1 take care of the basic quantity to be optimized, plus one isoperimetric condition. The terms in the constant Lagrange multipliers λ_2 and λ_3 cover the conditions on f(1-) and f'(1-) considered as constraints on F. The terms in the variable multipliers λ_4 and λ_5 cover the basic equation (3.4) and the basic inequality constraint (3.6) or (3.10). Specific assignments to λ_0 or λ_1 will be given later in treating specific optimization problems.

An essential distinction must be made in the single-bound case between the F' associated with the multiplier λ_3 and the F' associated with the multiplier λ_5 . The former F' includes the negative delta function associated with any negative jump in F, as the corresponding term in the integral of the λ_3 term over x is included in the relation which expresses b^2 as a functional of F or F'. The latter F' (associated with λ_5) does not include such a negative delta function, as relation (3.14) simply has no meaning at a negative jump in F. This distinction is important in setting transversality and corner conditions.

The derivative of G with respect to F' may be expressed

$$G_{F'} = 2\lambda_3/(1-x)^{1/2} - \lambda_5(x)(1-(A/D) + (2F'/D)).$$
 (3.17)

Since F(1) is unspecified, one transversality condition must be imposed at x = 1. Because

of the singularity in the λ_3 term in Eq. (3.17) this takes the form

$$\lim_{x \to 1^{-}} \left[\frac{2\lambda_3}{(1-x)^{1/2}} - \lambda_5 (1 - (A/D) + (2F'/D)) \right] = 0. \tag{3.18}$$

The transversality condition for F at x=0 appears only in the single-bound case $D^{-1}=0$ with $a^2 \neq 0$ and unspecified. This is not simply $G_{F'}=0$ there, because the integral of the λ_3 term in Eq. (3.16) includes the quantity $-2\lambda_3F(0+)$ from the jump in F at the origin; the quantity λ_3 considered as a function of x must be continuous across this initial jump. The transversality condition considered as a corner condition involves the jump in $G_{F'}$ and thereby excludes the λ_3 term. This distinction in treating the F' terms is that mentioned earlier. The transversality condition then appears as

$$\lambda_5(0) = 0 \quad \text{if} \quad a^2 \neq 0.$$
 (3.19)

The Euler conditions involve the parameter b and the variables z, f, and F. The first condition is simply $\lambda_3 b$ $\delta b = 0$, or, with b either zero or unspecified,

$$\lambda_3 = 0 \quad \text{if} \quad b^2 \neq 0. \tag{3.20}$$

Condition (3.18) immediately gives

$$\lambda_5(1) = 0 \quad \text{if} \quad b^2 \neq 0.$$
 (3.21)

The Euler equation for z is

$$G_{s} = -2z\lambda_{5} = 0. \tag{3.22}$$

This condition states that except at points for which $\lambda_5 = 0$ the condition z = 0 must be met; then F' must equal one of its specified bounds.

The Euler equation for f is

$$\lambda_4(x) = \lambda_0(1-x) + \lambda_1. \tag{3.23}$$

The integral of λ_4 which is a coefficient of F in G may be evaluated

$$2\int_{x}^{1} \frac{\lambda_{4}(\xi) d\xi}{(\xi - x)^{1/2}} = \frac{8\lambda_{0}}{3} (1 - x)^{3/2} + 4\lambda_{1} (1 - x)^{1/2}.$$
 (3.24)

The Euler equation for F is

$$\frac{d}{dx} \left[\frac{2\lambda_3}{(1-x)^{1/2}} - \lambda_5 \left(1 - \frac{A}{D} + \frac{2F'}{D} \right) \right] - \frac{2\lambda_2}{(1-x)^{1/2}} - 2 \int_x^1 \frac{\lambda_4 \, d\xi}{(\xi-x)^{1/2}} = 0. \tag{3.25}$$

Its first integral then gives us explicitly

$$\lambda_{\delta} \left(1 - \frac{A}{D} + \frac{2F'}{D} \right) = \frac{16\lambda_{0}}{15} (1 - x)^{5/2} + \frac{8\lambda_{1}}{3} (1 - x)^{3/2} + 4\lambda_{2} (1 - x)^{1/2} + 2\lambda_{3} (1 - x)^{-1/2}.$$
 (3.26)

Property (3.18) is satisfied, and conditions (3.20) and (3.21) are consistent. The right-hand side of (3.26) is termed the *switching function*. It is in the form of $(1-x)^{-1/2}$ times a cubic in x. Thereby it may have at most three zeros in the closed interval $0 \le x \le 1$. Zeros of the switching function (and thereby of λ_5) in this interval are termed switching points.

The corner condition in its conventional form applies only in the double-bound case, the case for which F must be continuous. This condition in this case is that $\lambda_5 = 0$ where F' is discontinuous. This information is supported by that in condition (3.22), which indicates that discontinuities in F' may be permitted. The variable z at such a point is simply undefined but not required to be zero.

In the single-bound case F may have negative jumps. In this case G is linear in F', the corresponding terms having the coefficient $2\lambda_3(1-x)^{-1/2}-\lambda_5(x)$. As mentioned before, the Dirac delta function in F' associated with a discontinuity in F is considered to be included in the λ_3 term but not in the λ_5 term. The increment of the integral I under a variation then includes a term $\delta(\lambda_5[F])$, where [F] is the jump in F at a point $x=x_0$. This increment in I is of the form λ_5 $\delta[F]$ for a variation in which x_0 does not change. This gives the corner condition $\lambda_5=0$, in accord with condition (3.22). For a variation in which x_0 changes the complete increment in I includes a term in [F] δx_0 times $\lambda_5' + \{\lambda_3(1-x)^{-3/2} - \lambda_5' - 2\lambda_2(1-x)^{-1/2} - 2\int_x^1 \lambda_4(\xi-x)^{-1/2} d\xi\}$. The curly bracket comprises the terms in the Euler equation for F (see Eq. (3.25)), and is thus zero. Hence, in the single-bound case, an interior $(0 > x_0 > 1)$ jump in F has associated with it a double-corner condition

$$\lambda_5(x_0) = 0, \quad \lambda_5'(x_0) = 0,$$
 (3.27)

which is essentially a condition on the switching function.

This double-corner condition is given by Leitmann [4] in discussing a control variable which may have an impulse (or delta function). Standard examples quoted involving an impulsive control variable do not have a scalar variable subject to a single bound, and involve a different type of condition on the Lagrange multipliers.

Most of the terms in G(x) from definition (3.16) are linear, and hence give no contribution to any higher-order variation. The nonlinear terms are quadratic and are $-\lambda_3 b^2 - \lambda_5 (D^{-1} F'^2 + z^2)$. We require I to be a minimum, and thereby require these terms to form a non-negative definite quadratic form. Thereby we obtain the sufficient second-variation requirements

$$\lambda_3 \le 0, \tag{3.28}$$

$$\lambda_{s}(x) \leq 0. \tag{3.29}$$

These conditions may also be obtained using a condition of the Weierstrass type [6]. Condition (3.29) is basic in our analysis, and, for example, condition (3.28) and the second half of condition (3.27) follow from it.

Condition (3.29) implies other conditions on the switching function λ_5 in the single-bound case. From condition (3.19) we have $\lambda_5'(0) \leq 0$ if $a^2 \neq 0$ and $\lambda_5(0) = 0$. From condition (3.21) we have $\lambda_5'(1-\epsilon) \geq 0$ for small ϵ if $b^2 \neq 0$ and $\lambda_5(1) = 0$; this is to be interpreted as requiring $\lambda_2 \leq 0$ in this case. And at an interior jump in F with condition (3.27) holding we have $\lambda_5''(x_0) \leq 0$. Since condition (3.29) can be readily checked, these other conditions are not needed.

4. Solution types and the limit $D^{-1} \to 0$. It is convenient to label the principal types of solutions. In the single-bound case $(D^{-1} = 0)$, if there is an interior jump in F, the switching function satisfies condition (3.27) there and thus has a double zero there. The function F can have no other interior jump, as this would require the switching function to have two double zeros; this is impossible with the cubic-type form of Eq.

(3.26). The switching function may have a single zero at x = 1 or one at x = 0, but not both.

A solution with an interior negative jump in F at x_0 and with $\lambda_5(1) = 0$ is termed of type A. In a general solution of this type $b^2 \neq 0$ from condition (3.21), and $\lambda_3 = 0$ from condition (3.20); with this type $a^2 = 0$ from condition (3.19). A solution with an interior negative jump in F at x_0 and with $\lambda_5(0) = 0$ is termed of type B. In a general solution of this type $a^2 \neq 0$ from condition (3.19); with this type $b^2 = 0$ from condition (3.21). A solution with an interior negative jump in F and with $a^2 = 0$, $b^2 = 0$ and $\lambda_5(0) \neq 0$, $\lambda_5(1) \neq 0$ is termed of free type. A solution with no interior jump in F and with $a^2 \neq 0$, $b^2 \neq 0$ and $\lambda_5(0) = \lambda_5(1) = 0$ is termed of open type. Solutions with only one zero of λ_5 in the range $0 \leq x \leq 1$ cannot satisfy the required conditions on f(1-) and f'(1-), and thus do not exist.

The term "free" was chosen for solutions of the free type because this is the type of solution which appears in the optimum lift problem which is free of an isoperimetric condition on the center of pressure. The solution is also free of singularities at the end points. The term "open" was chosen for solutions of the open type because the F-function has no jumps in the open interval 0 < x < 1.

In the double-bound case F is continuous and F' is either A or -D except at the switching points. With three switching points and F' following the sequence (A, -D, A, -D) a solution is termed of type A. With F' following the sequence (-D, A, -D, A) a solution is termed of type B. With two switching points and F' following the sequence (A, -D, A) the solution is termed of free type. With F' following the sequence (-D, A, -D) the solution is termed of open type. Solutions with only one switching point cannot satisfy the conditions (3.5) and (3.12) and thus do not exist.

In both the single-bound and double-bound cases solutions of the free type or of the open type may be considered as limiting cases of the solutions of types A and B. Hence most of our analysis is on types A and B; the other two types are special and are simpler.

The choice of terms for the solution types in the double-bound case was made so that the solutions would approach those of corresponding single-bound types in the limit $D^{-1} \to 0$. We turn next to a consideration of this limiting process.

The multiplier 1 - A/D + 2F'/D on λ_5 in the left-hand side of (3.26) equals (1 + A/D) if F' = A and equals -(1 + A/D) if F' = -D. Since $\lambda_5 \le 0$ from condition (3.29), F' = A if the right-hand side of Eq. (3.26) (the switching function) is negative and F' = -D if the right-hand side is positive.

In the limit $D^{-1} \to 0$ the intervals over which F' = -D must vanish, becoming the negative jumps in F of the single-bound case. In a solution of type B or of open type this limit gives the value $F(0) = -a^2$ at the origin of the corresponding single-bound type. In a solution of type A or of open type this limit gives the corresponding negative infinite slope of F for x = 1+; it gives no jump in F at x = 1. In a solution of type A or B or of free type the internal segment with F' = -D becomes the internal negative jump in F of the corresponding single-bound type. Corresponding solutions of type B for F and λ_5 are sketched in Fig. 1.

Of interest is the fact that the double zero of λ_{δ} corresponding to condition (3.27) for an internal jump appears naturally as the coalescence of two simple zeros in the double-bound case. This process is evident in Fig. 1. The derivation of condition (3.27) given above was found after the condition had been determined by arguments based upon condition (3.29) and upon the coalescence process described.

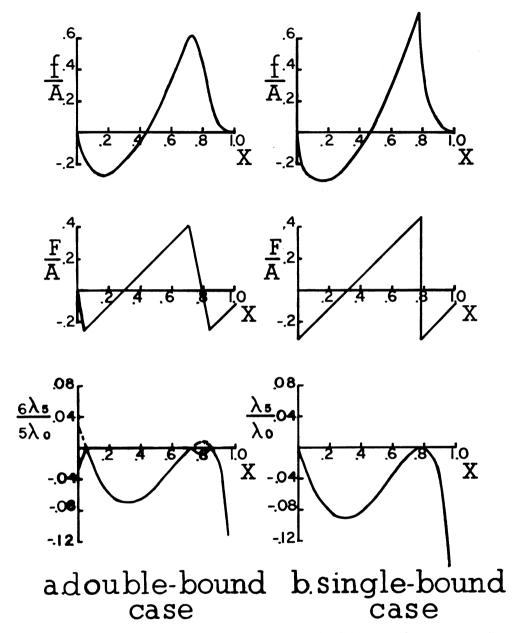


Fig. 1. Comparison of type B solutions in the double-bound and single-bound cases (D/A = 5).

Standard problems in the calculus of variations with inequality constraints are of the double-bound type. Our single-bound problems are unconventional not only in the transformation (3.15) applied to separate F(x) as a factor but also in the need to handle discontinuities in F. Consideration of the double-bound case was found to be instructive and to serve as a guide to the proper treatment of the single-bound case. It is reported here for these reasons, and because it has some potential direct application. It has served

its purpose at this point and is not considered further. Henceforth we treat only the single-bound case, and set $D^{-1} = 0$ in all our equations. The factor 1 - A/D + 2F'/D is equal to one in Eq. (3.26) for the switching function.

5. General solutions for optimization problems. We turn now to the specific solutions obtained for the functions F, f, and λ_{δ} . The interpretation in terms of physical optimization problems is postponed to later sections. There in specific problems values will be assigned to one or to both of the Lagrange multipliers λ_0 and λ_1 .

The solution for F in the interval $0 < x \le 1$ may be expressed

$$F = -a^2 + Ax$$
 for $0 < x < x_0$,
 $F = -a^2 + Ax - h$ for $x_0 < x \le 1$, (5.1)

where -h is the internal jump in F at $x = x_0$. Application of the transformation (3.4) gives

$$f = -4a^2x^{1/2} + (8A/3)x^{3/2} - 4h(x - x_0)^{1/2}, (5.2)$$

with the convention that the last term is dropped when it is imaginary, for $x < x_0$. This convention is used throughout this paper.

Condition (3.5) gives $f(1) = -4a^2 + (8A/3) - 4h(1 - x_0)^{1/2} = 0$, or $a^2 = (2A/3) - h(1 - x_0)^{1/2}$. Direct calculation gives

$$f'(1-) = b^2 = (8A/3) - 2hx_0(1-x_0)^{-1/2}, (5.4)$$

$$f''(1-) = (8A/3) + hx_0(2-x_0)(1-x_0)^{-3/2} \ge 0.$$
 (5.5)

The last result shows that condition (3.9) is always satisfied.

Direct computation also gives

$$S(x) = \int_0^x f \, dx = -\frac{8a^2}{3} x^{3/2} + \frac{16A}{15} x^{5/2} - \frac{8h}{3} (x - x_0)^{3/2}, \tag{5.6}$$

$$W = S(1) = -(32A/45) + (8h/3)x_0(1 - x_0)^{1/2}, (5.7)$$

$$V = \int_0^1 (1 - x) f \, dx = -(128A/315) + (16h/15)x_0(2 - x_0)(1 - x_0)^{1/2}, \tag{5.8}$$

$$M = \int_0^1 x f \, dx = W x_c = -(32A/105) + (8h/15)x_0(1 + 2x_0)(1 - x_0)^{1/2}. \tag{5.9}$$

The relation W = V + M may be checked.

The behavior of F for x > 1 is of minor interest as long as the basic inequality (3.6) is satisfied. The function F may be calculated from Eq. (3.3) with f'(x) = 0 for x < 1. For x - 1 small it may be expressed

$$F = F(1) - 2f'(1-)(x-1)^{1/2} + A(x-1) - \frac{4}{3}f''(1-)(x-1)^{3/2} + O((x-1)^2), \quad (5.10)$$
 with

$$F(1) = \frac{1}{3}A - h[1 - (1 - x_0)^{1/2}]. \tag{5.11}$$

Its asymptotic behavior for large x is

$$F \sim -\frac{1}{2}W(x-1)^{-3/2} + \frac{3}{4}V(x-1)^{-5/2} + 0((x-1)^{-7/2}).$$
 (5.12)

This information is generally sufficient to determine the qualitative behavior of F. We repeat here Eq. (3.26) for the switching function (now $\lambda_5(x)$) and express its derivatives

$$\lambda_5(x) = \frac{16\lambda_0}{15} (1-x)^{5/2} + \frac{8\lambda_1}{3} (1-x)^{3/2} + 4\lambda_2 (1-x)^{1/2} + 2\lambda_3 (1-x)^{-1/2}, \quad (3.26)$$

$$-\lambda_{\delta}'(x) = \frac{8\lambda_{0}}{3} (1-x)^{3/2} + 4\lambda_{1}(1-x)^{1/2} + 2\lambda_{2}(1-x)^{-1/2} - \lambda_{3}(1-x)^{-2/2}, \qquad (5.13)$$

$$\lambda_5''(x) = 4\lambda_0(1-x)^{1/2} + 2\lambda_1(1-x)^{-1/2} - \lambda_2(1-x)^{-3/2} + \frac{3}{2}\lambda_3(1-x)^{-5/2}.$$
 (5.14)

With an internal jump in F at $x=x_0$, condition (3.27) gives two equations for the four constant λ 's. Upon eliminating λ_0 , λ_1 , λ_2 , and λ_3 in turn we obtain four alternative relations, expressible as

$$\begin{vmatrix}
0 & 1 & 4 & 3 \\
1 & 0 & -1 & -1 \\
4 & 1 & 0 & -1 \\
3 & 1 & 1 & 0
\end{vmatrix}
\begin{vmatrix}
\frac{4}{15}\lambda_0(1 - x_0)^3 \\
\frac{4}{3}\lambda_1(1 - x_0)^2 \\
\lambda_2(1 - x_0)
\end{vmatrix} = 0.$$
(5.15)

These relations are used in evaluating the λ 's.

The values of the parameters in the various cases are presented in Table I. Switching functions are presented in terms of λ_1 in the table. They may be converted to forms in terms of λ_0 through the relation

$$\lambda_1 = -\frac{4}{5}\lambda_0(1-x_{\bullet}) \tag{5.16}$$

in the case of Type A, and through

$$\lambda_1 = -\frac{2}{5}\lambda_0(3 - 2x_0) \tag{5.17}$$

in the case of Type B. The condition on the λ 's characterizing type A is $\lambda_3 = 0(3.20)$, and characterizing type B is

$$(8\lambda_0/15) + (4\lambda_1/3) + 2\lambda_2 + \lambda_3 = 0, (5.18)$$

from (3.19).

In the case of solutions of both the free type and the open type the form of the switching function in terms of λ_1 alone (or of λ_0 alone) is not unique. The forms given in the table are for the particular case $\lambda_0 = 0$. The more general forms are

$$3(1-x)^{1/2}\lambda_{5}/8 = ((2-3x)^{2}/9)[\lambda_{1} + (2/15)(5-3x)\lambda_{0}]$$
 (5.19)

for solutions of free type, and

$$3(1-x)^{1/2}\lambda_{5}/8 = -x(1-x)[\lambda_{1} + \frac{2}{5}(2-x)\lambda_{\bullet}]$$
 (5.20)

for solutions of open type. The inequality (3.29) is satisfied if both $\lambda_1 + \frac{2}{3}\lambda_0 \leq 0$ and $\lambda_1 + (4/15)\lambda_0 \leq 0$ for a solution of free type, and if both $\lambda_1 + \frac{4}{5}\lambda_0 \geq 0$ and $\lambda_1 + \frac{2}{5}\lambda_0 \geq 0$ for a solution of open type.

6. Lift optimization problems. We consider two types of lift optimization problems, those in which no isoperimetric constraints are imposed beyond those universally set

Solutions to optimization problems				
Туре	A	В	free	open
$\frac{3a^2}{2A}$	0	$\frac{3x_0-2}{2}$	0	1
$\frac{3b}{4A}$	$\frac{2-3x_0}{1-x_0}$	0	0	2
$rac{3h}{2A}$	$\frac{1}{(1-x_0)^{1/2}}$	$\frac{2(1-x_0)^{1/2}}{x_0}$	31/2	0
x_{0}	$0 < x_0 < \frac{2}{3}$	$\frac{2}{3} < x_0 < 1$	$\frac{2}{3}$	_
$\frac{45W}{16A}$	$5x_0 - 2$	$2(4-5x_0)$	<u>4</u> 3	-2
$\frac{315 V}{32 A}$	$\left[-7x_0^2 + 14x_0 - 4 \right]$	$[2(12-21x_0+7x_0^2)]$	$\frac{20}{9}$	-4
$\frac{315M}{16A}$	$14x_0^2 + 7x_0 - 6$	$2(4+7x_0-14x_0^2)$	$\frac{44}{9}$	-6
x_{ϵ}	$\frac{14x_0^2 + 7x_0 - 6}{7(5x_0 - 2)}$	$\frac{4+7x_0-14x_0^2}{7(4-5x_0)}$	$\frac{11}{21}$	$\frac{3}{7}$
λ cond.	$\lambda_3 = 0$	Eq. (5.18)	$(\lambda_0 = 0)$	$\lambda_3 = 0 (\lambda_0 = 0)$
$\frac{3(1-x)^{1/2}\lambda_5}{8\lambda_1}$	$-\frac{(1-x)(x-x_0)^2}{2(1-x_0)}$	$\frac{x(x-x_0)^2}{(3-2x_0)}$	$\frac{(2-3x)^2}{9}$	-x(1-x)
Inequalities	λ_0 , $-\lambda_1$, $\lambda_2 < 0$	$-\lambda_0$, λ_1 ,	λ_1 , $-\lambda_2$,	$-\lambda_{1}$, λ_{2} $<$ 0
		$-\lambda_{2}$, $\lambda_{3} < 0$	$\lambda_{3} < 0$	

TABLE 1
Solutions to optimization problems

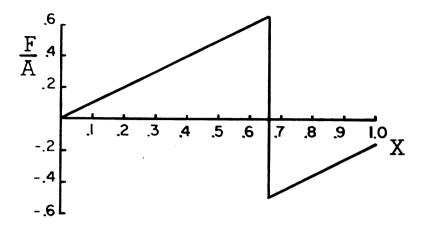
(such as fixed effective length), and those in which the load center variable x_{\bullet} is specified. For problems of the first type we set $\lambda_0 = 0$. If W is to be a maximum $\lambda_1 = -1$; if W is to be a minimum $\lambda_1 = 1$. The solution for maximum W is the free solution and is the one of normal interest in the sonic boom problem. The solution for minimum W (or maximum -W) is the open solution, and corresponds to an optimum bangless boom *upward* from an aircraft in steady flight; this solution has but peripheral interest.

The loading distribution for the normal lift problem free of additional isoperimetric constraint is shown in Fig. 2. In this solution f is never negative. The maximum effective gross weight in dimensionless terms is given by

$$W^* = 64A^*/135, (6.1)$$

with the asterisks from the reduction (2.16) replaced. This result is expressed in dimensional terms in the final section.

In the problem with an isoperimetric constraint on the variable x_e , the first two terms in expression (3.16) for G are $\{\lambda_0(x_e-x)-1\}f$ with W to be maximized. Thus



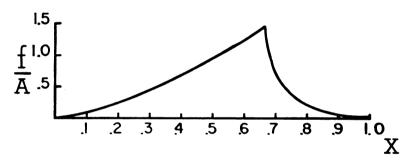


Fig. 2. Loading distribution and F-function for the normal lift problem (solution of free type).

 λ_0 is an undetermined Lagrange constant multiplier, while

$$\lambda_1 = -1 - \lambda_0 (1 - x_c). \tag{6.2}$$

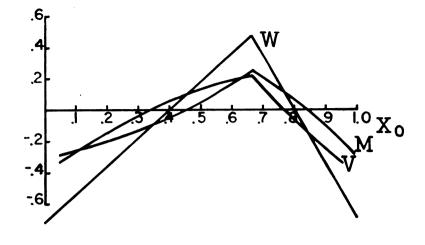
If the solution is of type A (5.16) also holds, and by using the expression for x_{ϵ} in Table I an expression for λ_0 or λ_1 in terms of x_0 may be obtained. In order that the inequality $\lambda_1 > 0$ or $\lambda_0 < 0$ be satisfied it is necessary that $5x_0 - 2 > 0$. Thus, for type A, we must have $\frac{2}{5} < x_0 < \frac{2}{3}$ and $11/21 < x_{\epsilon} < \infty$.

If the solution is of type B (5.17) also holds. The same argument may be followed, and we conclude that we must have $\frac{2}{3} < x_0 < \frac{4}{5}$ and $-\infty < x_c < 11/21$.

The solutions for maximum -W are obtained with $\lambda_1 = 1 - \lambda_0(1 - x_c)$ and are such that $0 < x_0 < \frac{2}{5}$ for type A and that $\frac{4}{5} < x_0 < 1$ for type B.

All of these solutions are shown in Fig. 3, which shows W, V, and M as functions of x_0 and W as a function of x_c . The latter plot represents an envelope of attainable values of W and x_c .

7. Volume optimization problems. We consider two volume optimization problems. The first is rather artificial, and demands that the volume V of a body of revolution be maximized between x = 0 and x = 1 (or L) with no constraint on the final area



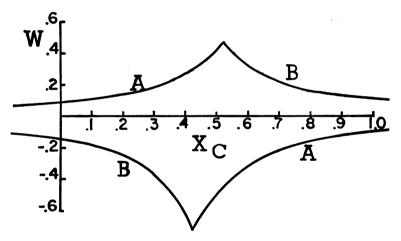
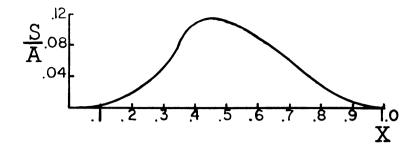


Fig. 3. The quantities W, V, and M as functions of x_0 , and W as a function of $x_c(A = 1)$.

S(1) = W. The body is thereby actually semi-infinite. The solution is obtained by setting $\lambda_1 = 0$ and $\lambda_0 = -1$. The solution is the one of free type, and is represented in Fig. 2 by interpreting f(x) as S'(x). Thus the solution is no different from that of the W-maximum lift problem with respect to the functions F and f = S'. There is a difference in the switching functions; In this case λ_5 is not that given in Table I but is the function obtained from (5.20) by setting $\lambda_1 = 0$.

The second volume problem considered is more realistic. It is required to maximize the volume V subject to the isoperimetric constraint of closure of the body S(1) = W = 0. We set $\lambda_0 = -1$ and let λ_1 be a Lagrange multiplier. From the basic inequalities the solution must be of type A and the requirement W = 0 dictates that $x_0 = \frac{2}{5}$. From (5.16), $\lambda_1 = 12/25$. The solution for $S(x) = \int_0^x f \, dx$ and $r_b(x) = (S/\pi)^{1/2}$ are presented in Fig. 4.

Other volume optimization problems may be posed, for example by specifying the value of S(1) = W as a constraint. These problems will yield some of the same solutions



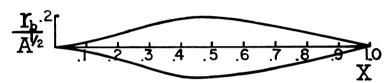


Fig. 4. Volume and radius distributions for optimum closed body of revolution.

obtained for lift problems. The physical requirement that S(x) > 0 over the range would exclude all solutions of type B and also solutions of type A with $0 < x_0 < \frac{2}{5}$.

8. Comments and conclusions. Certain results obtained in this particular problem in the calculus of variations may be expected to apply in a wider class of analogous problems. The control variable subjected to only an upper bound and which may have negative delta functions at singular points is associated with a one-signed switching function obeying a double-corner condition (3.28) at each singular point. A linear nonlocal relation between two variables imposed as an isoperimetric constraint is handled easily if the relation is in the form of a nonsingular integral equation; in this process the transposed integral transform of the Lagrange variable multiplier appears in the Euler equations. An approach using the Hamiltonian method of Pontryagin instead of the slack-variable method should yield essentially the same results.

The problem originally posed was one on the physical phenomenon of sonic boom. The main result is (6.1), giving the maximum effective gross weight attainable with incipient formation of a shock wave at the ground. To put this result into dimensional terms we reexpress Eq. (6.1) in terms of the dimensional variables W and A. The result is

$$W = \frac{64\rho_a a_a^2 M^2 L^{5/2} A}{135\beta}. (8.1)$$

For an isothermal atmosphere we obtain A from (2.8) and (2.11); this is repeated

$$A = \frac{2\gamma^{1/2}\beta^{3/2}g^{1/2}(\cos\phi)^{1/2}}{(\gamma+1)\pi^{1/2}M^4a_a \operatorname{erf}(\zeta)},$$
(8.2)

with

$$\zeta = (\gamma g h / 2a_a^2)^{1/2} \tag{8.3}$$

and h the flight altitude. Thus W is expressed as

$$W = \frac{128\gamma^{1/2}\beta^{1/2}g^{1/2}\rho_a a_a L^{5/2}(\cos\phi)^{1/2}}{135(\gamma+1)\pi^{1/2}M^2 \text{ erf }(\zeta)}.$$
 (8.4)

If the effective weight W in the direction ϕ comes only from the actual gross weight W_0 they are related by

$$W = W_0 \cos \phi. \tag{8.5}$$

Thus the actual gross weight permitted is proportional to $(\cos \phi)^{-1/2}$ in this case, which confirms that it is the ray directly beneath the flight path which is most critical.

The maximum pressure in the signal is proportional to $F_{\text{max}} = \frac{2}{3}AL$. Evaluated at the ground, this pressure p_{max} is that given by the direct theory of Sec. 2 times a reflection factor R which is usually 2. From (2.3) and (2.4) we obtain

$$\frac{p_{\text{max}}}{R} \left(\frac{\rho_a a_a}{\rho_a a_a} \right)^{1/2} = \frac{\rho_a a_a^2 M^2}{2^{1/2} \beta^{1/2} r^{1/2}} \cdot \frac{2AL}{3}. \tag{8.6}$$

Eliminating A through (8.1), we obtain

$$\frac{p_{\text{max}}}{R} \left(\frac{\rho_a a_a}{\rho_a a_a} \right)^{1/2} = \frac{45 \beta^{1/2} W (\cos \phi)^{1/2}}{2^{1/2} \cdot 32 L^{3/2} h^{1/2}} , \tag{8.7}$$

using the relation $h = r \cos \phi$. If (8.5) holds this pressure is proportional to $(\cos \phi)^{3/2}$, and is greatest immediately under the flight path.

The classic result in sonic-boom optimization theory is that of Jones [7] for the maximum gross weight attainable with a given shock strength Δp on the ground. In this theory the asymptotic N-wave shape is assumed. In the case corresponding to ours the F-function has a delta function of strength $W\beta/4\rho U^2L^{1/2}$ at x=0 and is zero over the open (dimensional) interval 0 < x < L. This gives a loading with $x_c = \frac{1}{3}L$. If L' is the half-thickness of the N-wave, the asymptotic theory gives $F_{\text{max}} = AL'$ and $W\beta/4\rho U^2L^{1/2} = \frac{1}{2}AL'^2$. Elimination of L', conversion of F_{max} into $\Delta p/R$, and substitution of A from (8.2) yields the result

$$W = \frac{2\pi^{1/2}(\gamma + 1)M^2hL^{1/2} \operatorname{erf}}{\gamma^{1/2}\beta^{3/2}g^{1/2}\rho_{\rho}a_{\rho}(\cos\phi)^{3/2}} \left(\frac{\Delta p}{R}\right)^2, \tag{8.8}$$

with Δp the allowable shock strength on the ground.

Some of these results, in particular forms of (8.4), (8.7), and (8.8), have been reported in [1] and [8], with $\cos \phi = 1$ and with erf (ζ) approximated by 1. The coefficients given in these references were incorrect, and should be replaced by the ones in this paper.

In the formulas given above a distinction is made between the speed of sound at the aircraft flight altitude and that at the ground. In an isothermal atmosphere, of course, there is no difference. The distinction is made so that the formulas may be applied approximately in a stratified atmosphere without winds which is not isothermal. In this case the altitude h should be reinterpreted as the streamtube width at the ground per unit azimuth angle at the aircraft.

The two formulas (8.4) and (8.8) give maximum gross weights which follow completely different scaling laws. The allowable gross weight following the Jones maximum (8.8) is relatively insensitive to Mach number and is but weakly dependent upon effective

length L. With Δp specified the allowable gross weight is proportional to the altitude h at which it flies.

The allowable gross weight following our bangless-boom optimum (8.4) is greatest at a low supersonic Mach number. It is highly sensitive (proportional to $L^{5/2}$) to effective length. And it is highly sensitive to aircraft flight altitude through the factor ρ_a , with the allowable gross weight increasing with decreasing altitude. A successful bangless-boom aircraft cannot fly too fast or too high, or be too short.

The formulas above have been based upon a signal in which the front and rear shocks form right at the ground. If the minimum rise time t_0 of the initial compression is specified, or, alternatively, if the maximum rate of pressure increase in the signature is specified as $p_{\rm max}/t_0$, (8.4) requires a minor revision. The value of W attainable is reduced by the factor $(1-3Ut_0/2L)$, which must be included in the right-hand side of (8.4). Eq. (8.7) is unchanged.

Other optimum sonic boom studies based upon a signal which is not an N-wave have appeared ([9], [10], and [11]). These have accepted the presence of a shock wave, and for the most part have concentrated on minimizing the strength of the bow shock.

More recently, Seebass and George [12], [13] have presented studies in which bow and tail shocks of equal strength are permitted in a signal which otherwise is limited by an upper bound on F'. The shapes obtained are heuristically derived, and their results reduce to ours in the case of specified zero shock strength in the free lift problem.

REFERENCES

- [1] W. D. Hayes, Sonic boom, Annual Review of Fluid Mechanics 3, 269-290 (1971)
- [2] F. E. McLean, Some nonasymptotic effects of the sonic boom of large airplanes, NASA Tech. Note TN D-2877 (1965)
- [3] W. D. Hayes, R. C. Haefeli, and H. E. Kulsrud, Sonic boom propagation in a stratified atmosphere, with computer program, NASA Contr. Rep CR-1299 (1969)
- [4] G. Leitmann, Variational problems with bounded control variables, in Optimization techniques with applications to aerospace systems (G. Leitmann, ed.), Academic Press, New York, 171-204 (1962)
- [5] A. Miele, Calculus of variations in applied aerodynamics and flight mechanics, in Optimization techniques with applications to aerospace systems (G. Leitmann, ed.), Academic Press, New York, 99-170 (1962)
- [6] F. B. Weiskopf, Jr., The optimization of certain aerodynamic parameters with the constraint of a bangless sonic boom, M.S.E. Thesis, Department of Aerospace and Mechanical Sciences, Princeton University (1971)
- [7] L. B. Jones, Lower bounds for sonic bungs in the far field, Aeron. Quart. 18, 1-21 (1967); see also J. Roy. Aeron. Soc. 65, 433-436 (1961)
- [8] W. D. Hayes, J. H. Gardner, D. A. Caughey, and F. B. Weiskopf, Jr., Theoretical problems related to sonic boom, in Third conference on sonic boom research, NASA SP-255, 27-31 (1971)
- [9] R. Seebass, Minimum sonic boom shock strengths and overpressures, Nature 221, 651-653 (1969)
- [10] A. R. George, Lower bounds for sonic booms in the midfield, A.I.A.A. J. 7, 1542-1545 (1969)
- [11] L. B. Jones, Lower bounds for the pressure jump of the bow shock of a supersonic transport, Aeron. Quart. 21, 1-17 (1970)
- [12] R. Seebass and A. R. George, Sonic boom minimization, Proc. Second Sonic Boom Symp., 1970, J. Acoust. Soc. Amer. 44, (1971)
- [13] A. R. George and R. Seebass, Sonic boom minimization including both front and rear shocks, AIAA J. (1971)