Appendix

The efficiency, which is defined as the ratio of power output P_0 to power input P_i, is taken to be proportional to Carnot efficiency, i.e.,

$$P_0/P_i = a(1 - x) \qquad 0 < a \leq 1 \tag{A1}$$

The parameter x is the ratio of radiator temperature T_r to input temperature T_i. The radiated power P_r, which is equal to the difference between P_0 and P_i, is given by

$$P_r = P_0 - P_i = bx^4 \tag{A2}$$

where $b = A_r\, \epsilon\sigma T_i^4$, with A_r being the radiator area, ϵ the relative emissivity, and σ the Stefan-Boltzmann constant. Eliminating P_i from Eqs. (A1) and (A2) gives

$$\psi = -ax^4(1 - x)/[a(1 - x) - 1] \tag{A3}$$

where $\psi = P_0/b$ is proportional to the power output per unit area of radiator. To maximize ψ, it is required that $d\psi/dx = 0$. This yields the following equation for the optimum value of x:

$$0 = 4a(1 - x)^2 - 5(1 - x) + 1$$

Thus for $a = 1$, the optimum value of x is 0.75, whereas in limit $a \to 0$, it is 0.80. This means that, even for a very poor Carnot engine ($a < 1$) that is cooled by radiation, the optimum radiator temperature is not much different from that of a perfect Carnot engine.

References

[1] Hirsch, R. L. and Holland, J. W., "Problems associated with the development of a thermionic conversion reactor," *Progress in Astronautics and Rocketry: Space Power Systems* (Academic Press, New York, 1961), Vol. 4, p. 389.
[2] Perry, L. W. and Pidd, R. W., "Nuclear reactor-thermionic space power systems," *Progress in Astronautics and Rocketry: Space Power Systems* (Academic Press, New York, 1961), Vol. 4, p. 399.
[3] Bernatowicz, D. T., "Thermionic converters in a large nuclear electric powerplant in space," ARS J. **32**, 1256–1261 (1962).
[4] Voorhees, B. G., "Progress toward nuclear thermionic space power," ARS Preprint 2577-62 (September 1962).
[5] Miller, B., "Nuclear thermionic power unit proposed," Aviation Week Space Technol. **75**, 61–62 (July 10, 1961).
[6] Wilson, V. C., private communication (December 1960).
[7] Ingold, J. H., "Calculation of the maximum efficiency of the thermionic converter," J. Appl. Phys. **31**, 769–772 (1961).

MAY 1963 AIAA JOURNAL VOL. 1, NO. 5

Nuclear Rocket Thrust Optimization Using Dynamic Programming

Robley E. George*
Douglas Aircraft Company Inc., Santa Monica, Calif.

The problem considered herein is that of maximizing the "burnout" velocity of a solid core nuclear rocket. The optimal thrust schedule has been shown, for a range of realistic parameter values, to differ from the well-known result for chemical systems because of the necessity to remove core fission product decay afterheat. This afterheat, which must be removed if the core is to remain intact, is a function of the reactor operating power schedule. The criterion of optimality is the maximization of the rocket final velocity under the constraints of a maximum and minimum allowable reactor power level corresponding to engine throttleable limits and a fixed initial propellant loading that is to be allocated either as useful thrust-producing propellant or as a core afterheat removal coolant yielding no useful thrust. The method of dynamic programming is used to perform the optimization. Numerical results are presented for an example problem that considers vertical, drag-free liftoff from a stationary earth with a uniform gravitational field. Means to relax all of these restrictions, along with other important extensions, are discussed. The effect of varying the problem parameters also is indicated, and implications thereof are pointed out.

Statement of the Problem

No doubt a lingering vestige of the chemical rocket propulsion era, it appears to be a common assumption that nuclear-powered rockets will perform at constant maximum power and thrust. In lieu of a better alternative, this certainly is a reasonable assumption. It is the purpose of this paper to examine that assumption and to determine whether or not a better alternative is available. The method of dynamic programming is well suited to answer just such a question.

Perhaps the definitive applications of the technique of dynamic programming to nuclear reactor systems optimization are those by Ash, Bellman, and Kalaba[1] and Kallay.[2, 3] The original works by Bellman, e.g., Refs. 4 and 5, remain the most readable, informative, and amusing introduction to the subject of dynamic programming. Further applications of dynamic programming to a wide range of optimization problems encountered in engineering are given by Kalaba.[6]

Received by ARS October 29, 1962. This paper was prepared under the sponsorship of the Douglas Aircraft Company Independent Research and Development Program, Account No. 81425-110, E.W.O. No. 51970. The writer counts it a pleasure to record that his stimulating introduction to matters programming, both Fortran and dynamic, was provided by N. Kallay. In a series of lectures given at the University of California at Los Angeles, R. E. Kalaba carefully nurtured both interests. His personal assistance in the preparation of this paper also is acknowledged gratefully.

* Design Engineer, Research and System Analysis, Space Propulsion Section, Missile and Space Systems Division.

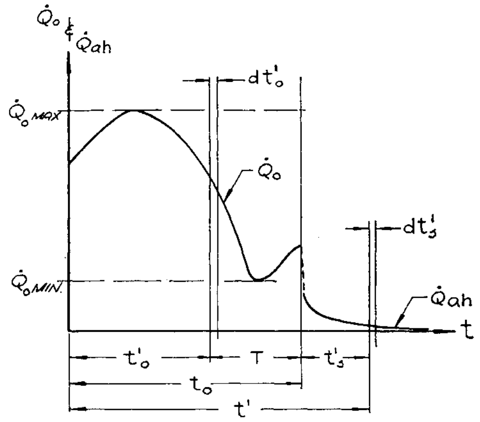

Fig. 1 Arbitrary power schedule and resultant afterheat power

Dynamic programming has been applied successfully to a number of rocket optimization problems. Typical, but by no means exhaustive, are the work by Ten Dyke[7] which treats the completely general multistage rocket problem and the work by Smith[8] which considers orbit transfer problems.

Pertinent to the problem considered in this paper are the following references. Berkovitz[9] establishes, among other things, that, under certain reasonable assumptions, the well-known optimal chemical rocket thrust schedule is to fire at constant maximum thrust. Fresdall and Babb[10] attempt to optimize, by a brute force enumeration technique, a reactor shutdown schedule in order to minimize Xe poisoning—a problem different from though related to the one considered herein.

Preliminary estimates of the amount of propellant required to cool convectively and hence to remove fission product decay afterheat from a solid core nuclear rocket reactor, if it is desired to preserve the core intact even temporarily, are of the order of 7% of total propellant mass. For not particularly ambitious nearby planetary travel, the afterheat removal propellant mass may be comparable to payload mass. Hence, it would appear that a significant percentage increase in payload mass might be achieved by power operation schedules that reduce afterheat generation. The ultimate (and only truly significant) objective of overall system optimization is hereby acknowledged but will not be considered in great detail here.

Consider first the vehicle dynamical equation

$$dV/dt = (F/M) - g \qquad V(0) = v \tag{1}$$

where

F = instantaneous rocket thrust
M = instantaneous rocket mass
g = gravitational acceleration (here assumed constant)
V = instantaneous rocket velocity
t = time

It will be noted that no drag force is included. As will be seen later in the dynamic programming formulation of the problem, the inclusion of a drag force either as an analytic expression or as numerical data would be a near trivial extension. Similar comments apply to varying g and a host of other "practical" extensions, as will be shown.

The author now will attempt to characterize the necessary afterheat removal coolant mass penalty resulting from a particular thrust schedule. One assumption would be to say that the afterheat removal coolant mass is proportional to the total afterheat energy generated by a particular thrust (and reactor power) schedule. Although the time rate of fission product decay energy is determined by a large number of radioactive fission fragments simultaneously decaying, each with its own respective half-life, the whole process may be characterized, with reasonable accuracy, by a relatively simple empirical equation.

Consider Fig. 1. Here $\dot{Q}_0$ is reactor operating power, ranging between corresponding engine throttleable limits, and $\dot{Q}_{ah}$ is fission product decay afterheat power, with various running times and time increments as indicated.

Glasstone (see Ref. 11, p. 119) gives, as the empirical afterheat power response at time t' due to a unit impulse of fission energy at time t_0',

$$h(t', t_0') = k_1(t' - t_0')^{-1.2} \qquad t' > t_0' \tag{2}$$

where $h(t', t_0')$ = unit impulse response, and k_1 = const. During a small time interval about time t_0', however, there is $\dot{Q}_0\,(t_0')\,dt_0'$ fission energy generated. Hence, the actual afterheat power at time t' due to the actual fission energy generated at t_0' is

$$\dot{Q}_{ah}\,(t', t_0') = h\,(t',t_0')\,\dot{Q}_0\,(t_0')\,dt_0' \tag{3}$$

which is to say that

$$\dot{Q}_{ah}\,(t',t_0') = k_1\,(t' - t_0')^{-1.2}\,\dot{Q}_0\,(t_0')\,dt_0' \tag{4}$$

Assuming that the empirical afterheat equation (2) is applicable for all time, one may integrate (4) to obtain the total afterheat energy resulting from the fission energy generated at t_0'. Thus

$$Q_{ah}(t_0') = \int_{t'=t_0}^{\infty} [k_1\,(t' - t_0')^{-1.2}\,\dot{Q}_0\,(t_0')\,dt_0']\,dt' \tag{5}$$

where $Q_{ah}\,(t_0')$ is the forementioned afterheat energy penalty. Performing the indicated integration, one obtains

$$Q_{ah}\,(t_0') = 5\,k_1\,\dot{Q}_0\,(t_0')\,(t_0 - t_0')^{-0.2}\,dt_0' \tag{6}$$

or

$$Q_{ah}\,(t_0') \cong k_2\,\dot{Q}_0\,(t_0')\,(t_0 - t_0')^{-0.2}\,\Delta t_0' \tag{7}$$

and, finally,

$$Q_{ah}\,(t_0') \cong k_2\,\dot{Q}_0\,(t_0')\Delta t_0'\,(T)^{-0.2} \tag{8}$$

where $k_2 = 5\,k_1$, and $T = (t_0 - t_0')$.

Of course, other characterizations of the afterheat penalty are conceived easily. For example, one might carry out the integration until such time as thermal radiation in space would equal or exceed the afterheat generation rate with the engine at some maximum allowable temperature. Other possibilities are immediate; however, this particular formulation is convenient from the computational standpoint of dynamic programming. The fact that the empirical afterheat equation may not be applicable over all time may be compensated for by judicious choice of the coefficient equating total afterheat energy and requisite coolant mass.

Using the original assumption concerning requisite afterheat removal coolant mass, one has

$$m_{ah}\,(t_0') = k_3\,\dot{Q}_0\,(t_0')\,\Delta t_0'\,(T)^{-0.2} \tag{9}$$

where $m_{ah}\,(t_0')$ is the coolant mass needed for afterheat removal due to fission energy generated at t_0', and k_3 is a constant reflecting fraction of afterheat to be removed convectively along with coolant specific heat, attainable exit temperature, etc.

For convenience, assume proportionality among reactor operating power, engine thrust, and propellant flow rate. Hence

$$F = k_4\,\dot{Q}_0 \tag{10}$$

and

$$dM/dt = -k_5\,\dot{Q}_0 \tag{11}$$

where k_i, i = 4,5 are constants greater than zero.

One thus is led to consider the following specific problem. Given a solid core nuclear-powered rocket with a fixed initial propellant loading and an engine design such that a portion of the total afterheat must be removed convectively using the necessarily remaining propellant as coolant, what is the best way to allocate the initial propellant, as a function of time, between thrust-producing propellant and afterheat removal coolant that produces no useful thrust so as to maximize vehicle burnout velocity?

One will see that the dynamic programming formulation of the problem will transform the foregoing two-point boundary value problem with undetermined end condition into a simple initial condition problem, the solution of which is simplicity itself.

Solution of the Problem

I. Analytical Aspects

Adopting the viewpoint of dynamic programming, define the function f to be the maximum velocity increment attainable by the nuclear rocket. The function f is, after all, what one is after! Clearly, f is a function of the initial rocket mass. Also, it depends upon the initial unallocated propellant mass. Precisely, define the following: $f\,(M, m) \triangleq$ the velocity increment attained by the nuclear rocket, starting with an initial rocket mass M and an initial unallocated propellant mass m, and following an optimal nuclear rocket reactor power control policy.

A comment concerning the relationship among the variables is, perhaps, in order. It is quite true that, at any time, M and m are related. Why, then, are the two variables indicated as being explicitly independent? The answer is that, although M and m are related, it is not known what that relationship is. Indeed, that is part of what is to be determined as a function of time. The situation is, mathematically, a non-Markovian process, as discussed, for example, by Bellman (see Ref. 5, pp. 4 and 54). To extricate oneself from the dilemma, one imbeds this particular problem in the whole class of problems involving all possible (or reasonable) relationships between M and m. In the nonce, the price of the extrication is a single added dimension to the problem.

Bellman's intuitively simple, yet powerful, principle of optimality[4] will be used which, when phrased in the context of the present problem, may be stated as follows: *An optimal nuclear rocket reactor power control policy has the property that, whatever the initial state (M, m) and initial decision q are, the remaining decisions must constitute an optimal policy with regard to the new state $(M - \Delta M, m - \Delta m)$ resulting from the first decision.*

Application of Bellman's principle of optimality leads to the following recursive formula for $f\,(M, m)$:

$$f\,(M, m) = \max_{\dot{Q}_{0_{\min}} \leq q \leq \dot{Q}_{0_{\max}}} \times \{\Delta v + f(M - \Delta M, m - \Delta m)\} \quad (12)$$

Note that in Eq. (12) Δv is the initial increment in velocity, and $f\,(M - \Delta M, m - \Delta m)$ is the best possible future increment in rocket velocity starting in the new state $(M - \Delta M, m - \Delta m)$. Both the initial return and the best possible future return are dependent upon the decision variable q, the power level at time T before shutdown. Hence, the maximum attainable total increment in velocity starting at the condition (M, m), i.e., $f\,(M, m)$, is the maximum of all the possible sums of initial velocity increment plus best possible future velocity increment starting in the new state resulting from the initial decision.

From (1) and (10), one has

$$\Delta v = [(k_4\, q/M) - g]\ \Delta t_0' - 0\ (\Delta t_0') \quad (13)$$

to terms that are of first order in $\Delta t_0'$.

From (11),

$$\Delta M = k_5\, q\ \Delta t_0' + 0\ (\Delta t_0') \quad (14)$$

Finally, from (14) and (9),

$$\Delta m = q\ (k_5 + k_3\ (T)^{-0.2})\ \Delta t_0' + 0\ (\Delta t_0') \quad (15)$$

Here q is any specific admissible value of $\dot{Q}_0$, which is a reasonable formulation since at any (M, m) one must decide upon some specific value of $\dot{Q}_0$.

The return function f is, however, explicitly a function only of the state variables M and m. Therefore, one must eliminate the time parameterization from the last three equations. For reasons best established by alternative attempts, it is convenient to consider the unallocated propellant mass m as the fundamental variable. At different amounts of the unallocated propellant mass, a new decision must be made as to how to allocate a portion of the remaining unallocated propellent. Hence,

$$\Delta v = \left(\frac{k_4 q - Mg}{Mq\ [k_5 + k_3(T)^{-0.2}]}\right) \Delta m + 0\ (\Delta m) \quad (16)$$

and

$$\Delta M = \left(\frac{k_5}{k_5 + k_3\ (T)^{-0.2}}\right) \Delta m + 0\ (\Delta m) \quad (17)$$

Substituting these relationships into (12) gives

$$f\,(M, m) = \max_q \left\{\left(\frac{k_4 q - Mg}{Mq\ [k_5 + k_3(T)^{-0.2}]}\right) \Delta m + f\left[M - \left(\frac{k_5}{k_5 + k_3\ (T)^{-0.2}}\right) \Delta m,\ m - \Delta m\right] + 0\ (\Delta m)\right\} \quad (18)$$

with bounds on the admissible range of q implied. By use of Taylor's theorem and (18), one finds

$$f\,(M, m) = \max_q \left\{\left(\frac{k_4 q - Mg}{Mq\ [k_5 + k_3(T)^{-0.2}]}\right) \Delta m + f\,(M, m) - \left(\frac{k_5}{k_5 + k_3\ (T)^{-0.2}}\right) f_M\ \Delta m - f_m\ \Delta m + 0\ (\Delta m)\right\} \quad (19)$$

where $f_x = \partial f/\partial x$.

Since the term $f(M, m)$ within the braces is unaffected by the maximization process (it already is the maximum burnout velocity), it may be removed and cancellation effected. Dividing the resulting equation by Δm and then letting Δm tend to zero yields the following limiting nonlinear partial differential equation that must be satisfied by the function $f(M, m)$:

$$0 = \max_q \left\{\frac{k_4 q - Mg}{Mq\ [k_5 + k_3(T)^{-0.2}]} - \left(\frac{k_5}{k_5 + k_3\ (T)^{-0.2}}\right) f_M - f_m\right\} \quad (20)$$

As an initial condition, one notes that

$$f\,(M, m = 0) \equiv 0 \quad (21)$$

which is to say that, at the condition any M and no unallocated propellant remaining, i.e., at burnout, the maximum attainable increment in velocity is identically zero.

Although only the first term on the right-hand side of (20) contains q explicitly, nevertheless, a q is implied in the terms involving the remaining duration of the process T. Although the appearance of T explicitly in (20) may not be aesthetic analytically, from the computational standpoint it presents no particular problem, as will be shown below. It remains a problem whether T may or may not be eliminated from (20), leaving it a function solely of the state variables (M, m) and the decision variable q. Indeed, it is a question whether the elimination would be desirable, since to get time histories one would have to go back over the initial solution and reparameterize.

Because of the somewhat involved nature of the limiting partial differential equation, in particular, the presence of T, the remaining duration of the firing period, it appears that a numerical or computational solution to the problem will be a necessity. Inasmuch as a numerical solution is necessary, it also is necessary then to have a numerical problem. Accordingly, a particular rocket was considered. The Appendix presents the details of this "example" nuclear rocket along with the derivation of estimates of its various significant parameters.

The analytical aspects of the problem now are left behind, and attention is directed toward how, precisely, calculations might be performed to determine the optimal power, thrust, and propellant flowrate schedules that maximize vehicle burnout velocity.

II. Computational Aspects

The basic equations, for computational purposes, are (18) and (21), repeated here for convenience:

$$f\,(M, m) = \max_q \left\{ \left(\frac{k_4 q - Mg}{Mq\,[k_5 + k_3(T)^{-0.2}]} \right) \Delta m + f\left[M - \left(\frac{k_5}{k_5 + k_3\,(T)^{-0.2}} \right) \Delta m,\; m - \Delta m \right] \right\} \tag{18a}$$

and

$$f\,(M, m = 0) = 0 \tag{21}$$

In order to facilitate conceptual understanding of the mechanics of obtaining the numerical solution to the problem, consider Fig. 2, which is simply the two-dimensional (state) space within which the state (vector) of the system is always defined.

The bounds of the state space are simply maximum values of interest for each of the state variables. For example, $M_{\max}$ and $m_{\max}$ would be, respectively, the initial mass and the initial propellant loading of the rocket. The point P_i represents the initial condition of the physical process, i.e., rocket initial takeoff weight and full tanks. The region A need not be considered, inasmuch as it represents the dry weight of the rocket. Region B likewise may be excluded since it corresponds to a situation wherein the unallocated propellant mass is greater than the total propellant mass aboard the rocket. The remaining area is the physically significant one. This region is quantized appropriately by a suitable grid.

Perhaps it should be mentioned here that the state space quantization need not be equal along each coordinate. Indeed, since, in general, the physical dimensions of each coordinate of the state space will be different, it actually is nonsensical to talk about equal or unequal quantization of the phase space coordinates.

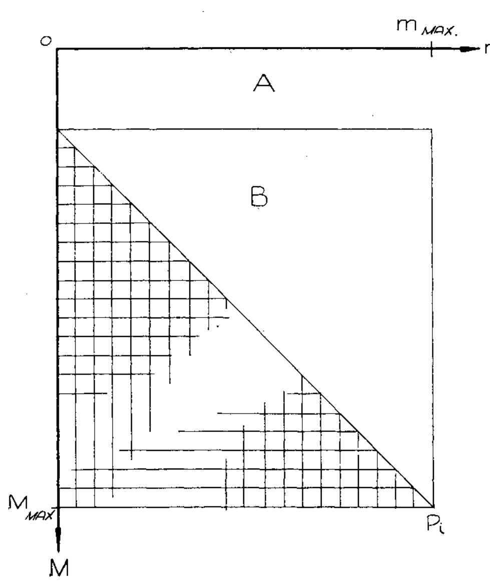

Fig. 2 Appropriate phase space

One starts the computation along the M axis ($m = 0$) and notes by (21) that, at *each* node point therein,

$$f\,(M, m = 0) = 0 \tag{22}$$

Along the M axis one has no decision to make concerning how to allocate m (since there is no m), and there is no time left in the process ($T = 0$). One thus associates with each node point on the M axis its value of f, viz., 0.

Next, one moves out the m coordinate to $m = \Delta m$. At this value of m, one evaluates the return function $f\,(M, m = \Delta m)$ for all values of M of interest by making use of (81a). Specifically,

$$f\,(M, m = \Delta m) = \max_q \left\{ \left(\frac{k_4 q - Mg}{Mq\,[k_5 + k_3(T)^{-0.2}]} \right) \Delta m + f\left[M - \left(\frac{k_5}{k_5 + k_3\,(T)^{-0.2}} \right) \Delta m,\; m - \Delta m = \Delta m - \Delta m = 0 \right] \right\} \tag{23}$$

But one already has evaluated f (any M, $m = 0$); in particular, its value is 0. Hence (23) reduces to

$$f\,(M, m = \Delta m) = \max_q \left\{ \left(\frac{k_4 q - Mg}{Mq\,[k_5 + k_3(T)^{-0.2}]} \right) \Delta m + 0 \right\} \tag{24}$$

As the whole of the problem calculations will be performed by digital computer machine computation, the discussion will be so oriented. The decision variable q is chosen successively out of the quantized admissible range of $\dot{Q}_0$, i.e., $\dot{Q}_{0\min} \leq q \leq \dot{Q}_{0\max}$. A convenient way to perform the maximization operation on a digital computer is simply to start with, say, the smallest q value, evaluate the expression in braces in (24), compare this with the value of the expression assumed using the next larger value of q, remember the better of the q's and the value it gives to the maximand of (24), choose the next larger q, evaluate, maximize (compare), store, and so on.

Having determined the maximum of the expression in the braces, one stores this as $f\,(M, m = \Delta m)$, and one also stores the maximizing q for that particular (M, m) location as $q_{\text{opt}}\,(M, m = \Delta m)$.

One then proceeds to a new M on the $m = \Delta m$ line and repeats the process. One thus associates (stores) with *each* node point on the $m = \Delta m$ line the value of $f\,(M, m = \Delta m)$ and the maximizing q.

Next one goes to the $m = 2\Delta m$ line and repeats the whole procedure. Specifically, one again uses (18a) to determine

$$(M, m = 2\Delta m) = \max_q \left\{ \left(\frac{k_4 q - Mg}{Mq\,[k_5 + k_3(T)^{-0.2}]} \right) \Delta m + f\left[M - \left(\frac{k_5}{k_5 + k_3\,(T)^{-0.2}} \right) \Delta m,\; m - \Delta m = 2\Delta m - \Delta m = \Delta m \right] \right\} \tag{25}$$

One sees that he already has f (any M, $m = \Delta m$), so that the maximization on the $m = 2\Delta m$ line is straightforward. It is true that interpolation may be necessary, but this is not a major difficulty. By this technique, one may "fill out" the region of interest in the (M, m) plane.

It is crucial to observe that these maximizations are single-stage decisions that can be made with utter ease. As stated, all one does is evaluate the sum of the initial velocity increment and the corresponding best future velocity increment for a particular q, compare this with the sum obtained by consideration of some other q, and do this for all possible q values, thus determining the maximum sum. The particular q that does the maximizing is then the optimal q. Remember also that, because of this method of computation, that is, starting out with the initial condition $f\ (M, m = 0) = 0$ and proceeding to larger values of m, for any particular value of m, say $m = m_1$, one already has evaluated $f\ (M, m - \Delta m = m_1 - \Delta m < m_1)$.

So far nothing has been said about the value of T to use in these calculations. Clearly, T is the time remaining in the allocation process starting at any (M, m) and proceeding optimally to the maximum attainable burnout velocity increment. Hence, at any $m = m_1 = n\Delta m$ line,

$$T = \Delta t + T' \tag{26}$$

where Δt is the duration of the immediate Δm allocation, and T' is the duration of the remaining optimal $(n - 1)\ \Delta m$ allocation process.

Obviously, on the particular line characterized by $m = \Delta m$, $T = \Delta t$. Therefore, (15) in the form

$$\Delta m/q = k_5 T + k_3(T)^{+0.8} \tag{27}$$

implicitly determines T, say,

$$T = T\ (M, m = \Delta m;\ q) \tag{28}$$

This determination, evidently, should precede the maximization operation conducted along the $m = \Delta m$ line. In anticipation of future need, one also should store, at each node point, that T corresponding to the maximizing q.

For all $m = n\Delta m$ lines, $n > 1$, a slight extension of the foregoing procedure will be necessary. At any phase space node and for all the q candidates, (26) will hold with T', in general being obtained by interpolation. One procedure for the determination of T might be as follows:

1) Make a first guess as to the value of T. This might be $T = T\ (M, m = n\Delta m) \cong T\ [M, m = (n - 1)\ \Delta m]$. Clearly the first guess ultimately would be increased.

2) Using (17) and the initial estimate of T, one could calculate ΔM.

3) Using this estimate of ΔM, one could interpolate to get an estimate of T'.

4) Using a slightly different form of (15), viz.,

$$\Delta m/q = k_5 \Delta t + k_3\ (T' + \Delta t)^{-0.2} \tag{29}$$

one could obtain an estimate of Δt.

5) A better estimate of T then would be $T_1 = T_0' + \Delta t_0$.

6) Iterating until $(T_{n+1} - T_n) < \delta_1$ or $(T_{n+1} - T_n)/T_n < \delta_2$, one would obtain T, say,

$$T = T\ (M, m = n\Delta m;\ q) \tag{30}$$

As discussed before, one would calculate the T's corresponding to two particular q candidates, effect the comparison, store the pertinent information for the better of the two, repeat for a new q candidate, and hence ultimately arrive at that $T = T\ (M, m)$ corresponding to the maximizing q for each (M, m) state. No doubt there are better techniques for the determination of T, but this is one method that works.

Note the effect that this constraint upon the admissible values of $\dot{Q}_0$, viz., $\dot{Q}_{0\min} \leq \dot{Q}_0 \leq \dot{Q}_{0\max}$, has on the facility with which one effects the maximization. Using classical methods, it compounds one's troubles no end, as is indicated, for example, in Ref. 9. Using the functional equation approach of dynamic programming, however, all such constraints are welcome; they reduce the work one must perform, as, intuitively, they should.

Note also that any extra information concerning the nature of the solution which, for some reason, one possesses is generally to no avail when applying classical variational techniques. Yet with dynamic programming one may use this information, as certainly seems appropriate. In this particular case, it has been noted that the required afterheat removal propellant mass is on the order of 7% of the total propellant mass. With this information, one need not consider the whole of the isosceles triangle defined by $M_{\max}$, $m_{\max}$ but only a small portion thereof by observing that changes in rocket mass M will go approximately as changes in unallocated propellant mass m. Obviously, considerations such as this could be programmed easily in a digital computer.

If, however, one were interested in a range of obtainable maximum velocity increments and desired to know the corresponding necessary initial rocket masses and propellant loadings—this often being called a parameter survey—one would consider the whole of the appropriate phase space. By so doing, one would obtain this parameter survey, each datum of information given being the best thing obtainable starting with the particular initial conditions and the whole of the parameter survey being obtained by *one* digital computer run, with corresponding economy reflected in computation cost and time.

A word about obtaining the particular solution corresponding to particular initial conditions might be of interest. Assuming that the rocket initial conditions indeed are represented by the point P_i in Fig. 2, by inspection of the computer results one immediately has the total firing time $T\ (P_i)$, the maximum attainable velocity $f\ (P_i)$, and the optimal reactor power level $\dot{Q}_{0\mathrm{opt}}(P_i)$ at which to fire initially. By Eq. (17), for example, one may proceed, successively, through the $m = n\Delta m$ lines, determining from the computer results at each line the optimizing $\dot{Q}_0$ and the remaining time of firing. Of course, this too could be done by the computer in the single run. The time history of the optimal reactor power thus is immediate, as are optimum thrust level and flow rate, since these all have been assumed proportional.

The foregoing comments are applicable to dynamic programming solutions in general, and it will be found that the functional equation formulation of dynamic programming will become a most useful tool in the analysis of rocket optimization problems.

Results

It is announced that the optimal, albeit unspectacular, "example" nuclear rocket power and thrust schedule for vertical flight from a stationary earth is constant "full blast" until such time as all remaining propellant must be used to remove the resultant afterheat. At this time, firing ceases immediately. Actually, there is possibly some power and thrust decay taking place in the final second or so of operation which, because of the particular phase space quantization used in this study, did not show up. More will be said of the fineness of the grid later; for the moment it merely is observed that, if this tailoff takes place within just the final few seconds, then one might as well forget about it for practical purposes.

Being rewarded for the effort with such undramatic results, one then sets out to determine interesting operating regimes in which the optimal power schedule diverged from the constant maximum level. One limiting condition is a zero-gravity field. It is rather clear that in this case the optimal thrust and reactor power schedules would be constant minimum (and this indeed was verified by numerical results), for this minimizes afterheat, and, since there is no penalty attached to an extended firing time, this then maximizes the amount of initial propellant that indeed may be fired as propellant. Of course, the burnout velocity for this case exceeded that for nonzero values of gravitational acceleration. What is important is that the optimal reactor power policy had changed.

Having now bounded the region of interest by forcing the solution to each of its constraints, the next logical step was to determine what conditions caused the changeover in optimal reactor power policy. This was done in a somewhat artificial way simply by performing the calculations for different values of gravitational acceleration.

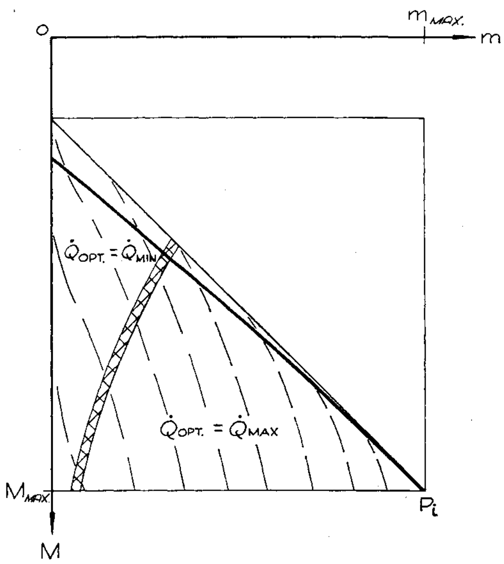

Fig. 3 Representation of numerical results

Anticipating perhaps some anxiety on the reader's part concerning computation times when considering a range of values of some of the parameters, it is appropriate to indicate the machine time required for these calculations. The rocket mass M—unallocated propellant mass m—state space was divided into a 51×51 grid, and the appropriate diagonal half of this space was considered. This gave something over $2500/2 = 1250$ node points at which all calculations were performed. Normalized reactor power was quantized to nine values ranging from 0.2 to 1.0. At each node point and for each q candidate, the appropriate remaining duration of the firing time was determined iteratively to within 1 sec. At each of the node points the optimal power level, maximum possible increment in rocket velocity, and correct remaining duration of the firing period were evaluated and printed. These calculations took approximately a couple of minutes on an IBM 7090 digital computer.

Of course, a necessary part of any analysis such as this must be a test of the sensitivity of the numerical results to the phase space quantization and, where appropriate, the decision variable quantization. In this particular problem, the acceptable tolerance in iterated remaining time also should be studied. A satisfactory, if not satisfying, method simply is to subdivide further until there is no significant change in the results. Inasmuch as the present study was solely of a preliminary investigation nature, no further refinement of the results was considered. With this general problem, viz., nuclear rocket afterheat, the appropriate refinement now is in the mathematical model, clarification of system mission, etc.

Returning to the results for various values of the gravitational acceleration, it is noted in passing that $g \sin\theta \leq g$, $0 \leq \theta \leq \pi/2$, so that in a very rough qualitative way a reduced value of g might be interpreted as a vehicle that has a component of thrust in the horizontal direction. This obviously is not correct, and, inasmuch as the correct way of doing it should be fairly simple—involving a few added dimensions to the problem—this observation is stated here only to indicate the area in which variable thrust will become optimal. Also, of course, vertical takeoff from, say, the moon is another situation in which a reduced value of g will obtain, although no blanket statement is appropriate, inasmuch as the nuclear rocket design may well differ for such a mission.

Briefly, whenever the physical parameters were such that a varying thrust was obtained, the optimal normalized reactor power started at the maximum rated value of 1.0 and remained there until a certain time (more accurately, certain values of the state variables) when it dropped fairly sharply to the minimum allowable level, at which it remained for the remainder of the firing time. In general, the descent of the optimal power level from maximum to minimum, in all cases studied, took place in one or two increments of unallocated propellant change. It was the position of this fairly abrupt power level reduction that varied, depending upon the particular values assigned to the parameters, that was the outstanding feature of the results. Of course, the burnout velocity also varied with the values of the parameters, but that is not the important aspect.

In general, the results obtained, for a given set of parameter values, were similar to those shown in Fig. 3. The single curved line in Fig. 3 represents the optimal trajectory through the phase space starting with the initial conditions characterized by the point P_i. From it one obtains the time histories of optimal power level, thrust, etc. As indicated, there is a region wherein optimal power level is maximum and another region in which optimal power level is minimum. The cross-hatched interface region is where optimal power level drops from maximum to minimum. It is this region that moves around, closer to, or away from the point P_i, depending upon the particular values of the parameters used.

Intuitive tendencies were borne out by numerical results. A high value of the afterheat coolant mass coefficient k_3 and a low value of g shifted the optimal policy transition region closer to the point P_i. The value of g was dominant in the location of the transition zone.

It would appear that a simple, near-optimal power schedule would be to fire at constant maximum rated power until a predetermined time at which power would be reduced to the minimum level and held there until shutdown. This certainly would be easy to effect, and it is amusing to note that in practice one actually would take some small amount of time to bring about the reduction in power. Thus, optimal policy could be approached to a high degree by a practical policy.

The family of dashed lines on Fig. 3 are isomaximum attainable burnout velocity curves and are to be interpreted as follows. The locus of points characterized by any one dashed line indicates all possible relationships between initial rocket mass and initial propellant mass, starting from which, and following an optimal reactor power control policy, one will achieve a certain final velocity increment, while at the same time providing for adequate cooling of the core during afterheat removal. The difference between initial rocket mass and initial propellant mass is, of course, the total dry weight of the rocket, and one sees that by this means one may vary that weight while accomplishing a specific mission, i.e., providing a specific velocity increment. The merit of such a parameter survey is self-evident.

It is remarked, perhaps trivially, that the general shape of the isoburnout velocity curves is as expected intuitively. For a unit increase in initial unallocated propellant mass (which is almost a unit increase in thrust producing propellant mass), one sees that the initial total rocket mass can be increased something more than this, and the same final burnout velocity still can be obtained. This certainly had better be the case, for adding a unit increment in initial propellant mass increases the total initial rocket mass by at least that amount.

The importance of this work, this writer feels, lies not in any particular results, but rather in the presentation of a logical, systematic way in which one, so interested, could go about determining optimal solid core nuclear rocket trajectories and power schedules.

Extensions are immediate. Clearly, one is not interested simply in vehicle velocity but also in the height at which it is attained. Atmospheric drag might be important and might depend upon both vehicle velocity and altitude. Curvilinear trajectories, angles of thrust, gravitational attraction depending upon altitude, possible utilization of afterheat coolant gas for thrust at, in general, a lower specific impulse, rocket thrust varying with altitude, and rocket specific impulse varying with power and thrust level are but a few of the refinements that, unquestionably, must be made if "practical" results are to be obtained. Yet, all of these necessary refinements are indeed simply extensions of the present formulation, which is not to say that the person effecting them would not deserve considerable credit.

The extensions will, in general, involve added dimensions to the problem. To surmount the high-speed memory limitations attendant such situations, the technique of polynominal approximation to the return function suggested by Bellman and Dreyfus[12] and employed by Aoki[13] appears attractive.

Appendix: Numerical Details

Herein the characteristics of the "example" rocket will be generated.

A. Nuclear Rocket Initial Conditions

$$M_{\max} = 300{,}000 \text{ lbm}$$

$$m_{\max} = 240{,}000 \text{ lbm}$$

and therefore

$$\lambda = 5$$

$$F_{\max} = 360{,}000 \qquad F_{\min} = 0.2\, F_{\max}$$

therefore

$$\left.\frac{F}{W}\right)_0 = 1.2$$

$$I_s = 800 \text{ sec} = \text{const}$$

therefore

$$\dot{M}_{\max} = 450 \text{ lbm/sec}$$

$$t_{b_{\min}} = 533 \text{ sec}$$

all of which neglects afterheat removal.

B. Afterheat Parameters

First a check on the applicability of the afterheat power response equation will be performed over all time from shutdown to infinity. Equation (5) may be expressed as

$$Q_{ah}(t_0') = \int_{t'=t_0}^{T^*} [k_1\,(t' - t_0')^{-1.2}\,\dot{Q}_0(t_0')\,\Delta t_0']dt' + \int_{t'=T^*}^{\infty} [k_1\,(t' - t_0')^{-1.2}\,\dot{Q}_0(t_0')\,\Delta t_0']dt' \quad \text{(A1)}$$

Integrating, one obtains

$$Q_{ah}(t_0') = K_1 \left\{ (t' - t_0')^{-0.2}\Big|_{t'=T^*}^{t_0} + (t' - t_0')^{-0.2}\Big|_{t'=\infty}^{T^*} \right\} \quad \text{(A2)}$$

with $K_1 > 0$.

Glasstone (see Ref. 11, p. 118) states that the afterheat response equation is applicable for something like a couple of weeks. Accordingly, we pick some representative values for the foregoing times and see how the two time functions compare in size. For the values $T^* = 10^6$ sec, $t_0 = 10^3$ sec, and $t_0' = 10^2$ sec, one finds

$$Q_{ah} = K_1\{(0.26 - 0.06) + (0.06 - 0)\} \quad \text{(A3)}$$

which says that all of the afterheat generated from $t' = T^* = 10^6$ sec on to infinity amounts to $(0.06/0.25) \approx \frac{1}{5}$ of the total amount of afterheat energy. Hence, one might presume that any of these results are "$\frac{4}{5}$" right, which is actually quite a respectable coefficient in engineering work. Also, of course, it must be that the afterheat equation slowly diverges from the physical facts rather than suddenly predicting outlandishly erroneous results. Hence, one concludes that these results are something better than "$\frac{4}{5}$" correct and leaves the matter, completely satisfied. (Clearly, we simply are trying to obtain some estimates of the error involved. Those interested may pursue the matter further.)

It remains to evaluate k_1 and, more important, k_3. Although the simple analytic form of Eq. (2), the Way and Wigner equation, will be used to characterize afterheat generation, the coefficient will be evaluated by means of the presumably more accurate equation of Untermyer and Weills.

Starting with Eq. (6),

$$Q_{ah}\,(t_0') = 5\,k_1\,\dot{Q}_0\,(t_0')\,(t_0 - t_0')^{-0.2}\,dt_0' \quad \text{(A4)}$$

one may assume, for the moment, a constant reactor power and integrate over all operating time to obtain total afterheat energy. So doing, one gets

$$Q_{ah} = (5\,k_1/0.8)\,(t_0)^{+0.8}\,\dot{Q}_0 \quad \text{(A5)}$$

which, with the reader's indulgence, may be written as

$$k_1 = (0.8/5)(Q_{ah}/\dot{Q}_0)\,[1/(t_0)^{+0.8}] \quad \text{(A6)}$$

Now referring to appropriate integrated Untermyer and Weills curves that are available from a number of sources, one easily may obtain an estimate of k_1. For example, taking $Q_0 = 10^6$ Btu/sec and $t_0 = 700$ sec, and from the curves noting that the total resultant afterheat energy generated from such a process would amount to about 2×10^7 Btu, one sees that

$$k_1 = \frac{0.8 \times 2 \times 10^7}{5 \times 10^6 \times (700)^{0.8}} = 0.017 \quad \text{(A7)}$$

with the units of k_1 being defined by the foregoing discussion.

As another check, let us see if the forementioned amount of afterheat energy is a reasonable amount considering the power level and firing time. Clearly the afterheat energy corresponds to 20 sec of full power operation. Therefore, the afterheat energy amounts to (20/720) of total energy generated, and this is about 2.8%. Now it is known that approximately 7% of total nuclear fission energy appears a "delayed" energy of some form. It seems as if this estimate gives a number about one half of this.

This brings up another point that is sure to rear its ugly head sooner or later. One observes that, *during* the firing period, it is true that some afterheat is being generated which increases just slightly the total power level of the reactor. One ramification of this is that the total afterheat energy, from reactor shutdown on, will be less than that indicated simply by all "delayed" energy. Furthermore, since afterheat is monotone decreasing with time, this discrepancy might be sizable, and indeed it appears to be so.

Another implication is that any analytical description that assumes reactor power level and fission power level identical is in error by precisely the forementioned amount. Accordingly, the results of the optimization are in error by some amount. The next question is, how much in error? One already has seen that the afterheat energy corresponds to a few percent of fission energy. Thus if one makes a small change in fission power level to account for the concomitant afterheat power, he will cause, at most, a second-order change in resultant afterheat and, presumably, optimal policy. So much for that!

With the foregoing estimate of k_1, one has, immediately, that

$$k_2 = 5\,k_1 = 0.085 \quad \text{(A8)}$$

Next it remains to establish the relationship between resultant afterheat and necessary afterheat removal coolant mass. For reasons not particularly germane to this discussion, it might be desirable to convectively cool and remove, say, two thirds of the total resultant afterheat energy. For similarly enigmatic reasons, it might be desirable to maintain a seemingly low exit coolant gas temperature. One may say that one might be satisfied with having the coolant gas exit at around 1400°R or, more specifically, that one could get 4.7 $\times$ 10^3 Btu of energy into each pound of afterheat removal coolant, on the average. Some other number might be more appropriate; what is needed here is *a* number.

Accordingly, one has

$$m_{ah} = \left(\frac{2}{3}\right) \frac{Q_{ah}}{\Delta Q_{ah}/\Delta m_{ah}} = \left(\frac{2}{3}\right) \frac{Q_{ah}}{4.7 \times 10^3} = 1.4 \times 10^{-4} Q_{ah} \quad \text{(A9)}$$

Combining (A4) and (A9), one obtains

$$m_{ah}(t_0') = 1.4 \times 10^{-4} \times 0.085 \times \dot{Q}_0 (t_0') (t_0 - t_0')^{-0.2} dt_0' \quad \text{(A10)}$$

which results in

$$k_3 = 1.2 \times 10^{-5} \quad \text{(A11)}$$

in the nomenclature of Eq. (9). Of course, for this specific value, k_3 has specific dimensions; these are defined implicitly in the discussion.

C. Thrust and Flow Rate Coefficients

From the previously stated I_s, which implies an exit gas temperature, and the maximum propellant flowrate, one obtains as a value for the maximum rated power level

$$\dot{Q}_{0_{\max}} = 6.55 \times 10^6 \text{ Btu/sec} \quad \text{(A12)}$$

$\dot{Q}_{0_{\min}}$ shall be equal to $0.2 \times \dot{Q}_{0_{\max}}$. With this power level and thrust level, one sees that

$$k_4 = 5.5 \times 10^{-2} \text{ lbf/(Btu/sec)} \quad \text{(A13)}$$

Similarly, using rated flow rate and reactor power, one obtains

$$k_5 = 6.9 \times 10^{-5} \text{ (lbm/sec)/(Btu/sec)} \quad \text{(A14)}$$

These k_i are as defined in Eqs. (10) and (11).

D. Numerical Comparison between Optimal Policy and Constant Maximum Power Policy

Illustrative of the advantages to be accrued by proceeding optimally, a brief comparison will be made between the final velocity of the example rocket when following an optimal reactor power schedule and when simply firing at constant maximum rate.

For certain particular values of the significant coefficients, a maximum attainable vehicle velocity of 31,131 fps was obtained starting at the initial conditions indicated by the point P_i. A total firing time of 713 sec corresponded to the optimal reactor power schedule. The results of this particular run just about correspond to the situation illustrated in Fig. 3.

Perhaps it should be pointed out that, although the drop in optimal power level from maximum to minimum for this particular case occurs at about one fifth of the way from completion of the firing process when viewed from the standpoint of the state variables considered, the drop also occurs about midway in the process when viewed as taking place in time. This is perhaps a more dramatic way of pointing out the fact that, if one wishes to proceed optimally, he is really going to have to do something about it.

Using the same particular values of the parameters, one may obtain, for the simple trajectory used, the burnout velocity of the rocket, assuming that it is fired at a constant maximum rate until such time that, because of the particular power schedule, all remaining propellant must be used for afterheat removal. As a first guess, one might assume the rocket to fire 455 sec. By the standard formula,

$$\Delta v = c \ln\lambda - gt_0 \quad \text{(A15)}$$

one obtains a burnout velocity of 29,655 fps. This corresponds to the following situation:

Propellant actually fired	=	6370 slugs
Vehicle dry mass	=	1870 slugs
Necessary afterheat coolant	=	1090 slugs
Total		9330 slugs

However, the initial takeoff mass was 9320 slugs. Accordingly, the initial guess of firing time corresponds to an impossible situation. One notes in passing that the optimally fired rocket betters by (31,131 − 29,655)/(29,655) × 100 = (1476/29,655) × 100 = 5% of the burnout velocity of the constant maximum fired rocket even when the latter performs the impossible.

References

[1] Ash, M., Bellman, R., and Kalaba, R., "On control of reactor shutdown involving minimal xenon poisoning," Nucl. Sci. Eng. **6**, 152–156 (1959).

[2] Kallay, N., "Dynamic programming and nuclear reactor systems design," Nucl. Sci. Eng. **6**, 315–325 (1960).

[3] Kallay, N., "An example of the application of dynamic programming to the design of optimal control programs," Inst. Radio Engrs. Trans. Auto. Control **AC-7**, 10–21 (1962).

[4] Bellman, R., *Dynamic Programming* (Princeton University Press, Princeton, N. J., 1957), p. 83.

[5] Bellman, R., *Adaptive Control Processes, A Guided Tour* (Princeton University Press, Princeton, N. J., 1961), pp. 4 and 54.

[6] Kalaba, R., *Introduction to Dynamic Programming* (Academic Press, New York, to be published).

[7] Ten Dyke, R. P., "Computation of rocket step weights to minimize initial gross weight," Jet Propulsion **28**, 338–340 (1958).

[8] Smith, F. T., "Optimization of multistage orbit transfer processes by dynamic programming," ARS J. **31**, 1553–1559 (1961).

[9] Berkovitz, L., "An optimal thrust control problem," J. Math. Anal. Appl. **3**, 122–132 (1961).

[10] Fresdall, J. R. and Babb, A. L., "Xenon-135 transients resulting from time-varying shutdown of thermal reactors," Trans. Am. Nucl. Soc. **4**, 316–317 (1961).

[11] Glasstone, S., *Principles of Nuclear Reactor Engineering* (D. Van Nostrand Co. Inc., Princeton, N. J., 1955), pp. 118 ff.

[12] Bellman, R. and Dreyfus, S., "Functional approximations and dynamic programming," Math. Tables and Other Aids to Computation **13**, 247–251 (1959).

[13] Aoki, M., "Dynamic programming and numerical experimentation as applied to adaptive control," Dissertation, Univ. Calif., Los Angeles, Calif. (November 1959).