Time-Optimal Control of Multivariable Systems Near the Origin

G. J. OLSDER

Communicated by J. V. Breakwell

Abstract. This article deals with bang-bang solutions of linear time-optimal control problems. Linear multivariable systems are considered which have one or more control components. It is shown in which way the control components act together to make the system achieve the ultimate aim (namely, the origin in the state space) as quickly as possible. The theory only applies to initial positions sufficiently near the origin. Criteria are given which give the number of switches per control component.

Asymptotic dependences of the switching times and the final time on the distance of the initial position from the origin are established. The theory provides a numerical procedure to calculate the time-optimal control. These calculations are very simple. Basic to the proof of these results is a generalized implicit function theorem due to Artin (Ref. 1).

Key Words. Bang–bang control, time-optimal control, switching surfaces, multidimensional control problems, control theory.

1. Introduction and Discussion of the Results

A system is described by the differential equation

$$\dot{x}(t) = Ax(t) + Bu(t),$$

(1)

where the dot denotes the derivative with respect to time. The $n$-vector $x$ with components $x_1, \ldots, x_n$ represents the state of the system; the $r$-vector $u = (u_1, \ldots, u_r)'$, the control, is a measurable vector-valued function of
time; the prime denotes the transpose. The matrices $A$ and $B$, of appropriate sizes, have constant elements. We assume that the columns of $B$ (denoted by $b_i$, $i = 1, ..., r$) are linearly independent. On account of the sizes of $A$ and $B$, the system (1) is called an $(n, r)$ system.

It is required to choose the control $u(t)$, whose components must satisfy

$$|u_i(t)| \leq 1, \quad i = 1, ..., r,$$

in such a way that the system is steered from the initial point $x = \epsilon x_0$ ($x_0$ is an $n$-vector with length one and $\epsilon$ is a positive scalar) at $t = 0$ to the origin in the state space as quickly as possible.

Under certain conditions concerning the system (it must belong to a subclass of controllable systems to be defined), the following facts about the time-optimal control (t.o.c.), which is bang–bang (Ref. 2), will be established for almost all directions $x_0$ and for $\epsilon$ sufficiently small.

(i) All control components together have precisely $n - 1$ switches. For the validity of this property, it is not necessary that all the eigenvalues of the matrix $A$ be real. It is known (Ref. 3) that, if these eigenvalues are real, each component separately has at most $n - 1$ switches. Also, criteria will be given which give the number of switches per control component.

(ii) If $\epsilon$ is supposed to be variable, then the switching times and the final time are analytic functions of $\sqrt[\varrho]{\epsilon}$, where the quantity $\varrho$ is the smallest natural number $\geq (n/r)$ ($x_0$ remains fixed and $\sqrt[\varrho]{\epsilon}$ is real). These functions have the value zero if their argument is zero; and, if they are developed in a Taylor series with respect to their argument, the coefficients of the linear terms are not equal to zero. Hence, for $\epsilon$ sufficiently small, the final time of the maneuver is approximately proportional to $\sqrt[\varrho]{\epsilon}$.

It follows that, for $\varrho > 1$ (i.e., $r < n$), we must know the switching times and the final time very exactly in order to steer the system from $\epsilon x_0$ to the origin. An error of order $\epsilon$ in one of the switching times or the final time (such an error is small with respect to the duration of the whole process if $\epsilon$ is sufficiently small) causes the final situation of the system to have again a distance of order $\epsilon$ from the origin. In this way, we can conclude that, for $\varrho > 1$ and $\epsilon$ sufficiently small, the criterion of time optimality is not a fortunate one, if we want to steer the system exactly to the origin. This is a mathematical formulation of the generally accepted engineering principle that, in order to reduce small initial perturbations, one should not use bang–bang control (which is moreover often difficult to realize for the final time sufficiently small). Instead of this, a continuous, linear feedback control is frequently used.
The essential idea of obtaining the results is the following. Let the \( n - 1 \) switching times be denoted by \( s_1, \ldots, s_{n-1} \) (only in this section) and the final time by \( T \). If the optimal control is substituted in (1), then, because \( x(T) = 0 \), \( n \) scalar equations result which can symbolically be written as

\[
 f_i(\epsilon, s_1, \ldots, s_{n-1}, T) = 0, \quad i = 1, \ldots, n. \tag{2}
\]

The functions \( f_i \) are analytic functions of their arguments in a neighborhood of \( \epsilon = s_1 = \cdots = s_{n-1} = T = 0 \). Because \( s_1, \ldots, s_{n-1} \), and \( T \) are analytic functions of \( \sqrt[2]{\epsilon} \), the corresponding series expansions can be substituted in (2), and that provides a method to determine the coefficients of these series. The actual calculation of these coefficients involves sets of algebraic, recurrence, and mostly linear equations. This provides a quick and simple numerical procedure to calculate the to.c. However, a serious drawback is that \( \epsilon \) must be sufficiently small.

Example 1.1. Consider \((4, 1), (4, 2), (4, 3), \) and \((4, 4)\) systems. The approximate dependence of the switching times and the final time is proportional to \( \sqrt[2]{\epsilon}, \sqrt[3]{\epsilon}, \sqrt[4]{\epsilon}, \) and \( \epsilon \), respectively. Hence, if we consider a \((4, 2)\) system, the addition of a new control component does not change the order with which, for instance, the final time changes with respect to sufficiently small \( \epsilon \).

Example 1.2. Consider a tractor with three carts, as shown in Fig. 1. Suppose that initially the combination of tractor and carts follows a rectilinear track (i.e., the middles of all eight axles are exactly on the track). The forward speed is constant. The tractor-driver is asked to steer the combination to another, parallel and nearby, rectilinear track in the least possible time (the middles of all axles must be on this new track). The steering-wheel is the control. It is assumed that no ditches or other obstacles exist in the neighborhood of the experiment. The distance between the two tracks is given by \( \epsilon \); and, for small \( \epsilon \), a linear vector differential equation suffices to describe the equations of motion of tractor and carts. In this example \( n = 8, r = 1 \), hence \( p = 8 \). For \( \epsilon \)

![Fig. 1. Tractor and carts.](image-url)
sufficiently small, the final time $T$ is analytic with respect to $\sqrt[3]{\varepsilon}$; and, considering only the first term in the Taylor expansion, $T = c \sqrt[3]{\varepsilon}$, $c$ being a constant. Hence, in order to cut down the final time by a factor 2, $\varepsilon$ must be cut down by a factor $2^8 = 256$.

Before the exact formulation of the properties mentioned will be given, the train of thought will be shown in the next section which is restricted to $(2, 1)$ systems and which is explicatory in character.

2. Elucidation by Means of $(2, 1)$ Systems

In this section, $n = 2, r = 1$, hence $p = 2$. The claim is that the t.o.c. has one switch, which will occur at time $t_{11}$. Suppose that the t.o.c. starts with a plus sign (the other case can be treated analogously). Equation (2) in this case yields:

$$0 = cx_0 + \left( \int_0^{t_{11}} - \int_{t_{11}}^T \right) \exp(-sA) b_1 \, ds. \quad (3)$$

If $\exp(-sA)$ in this equation is developed in its Taylor expansion, after which the integration can be carried out, the result is

$$0 = (\sqrt[3]{\varepsilon})^2 x_0 + b_1(2t_{11} - T) - (Ab_1/2!)(2t_{11}^2 - T^2)$$
$$+ (A^2b_1/3!)(2t_{11}^3 - T^3) \cdots. \quad (4)$$

Another claim is that convergent power series exist of the form

$$t_{11} = \sum_{i=1}^{\infty} \alpha_i(\sqrt[3]{\varepsilon})^i, \quad T = \sum_{i=1}^{\infty} \beta_i(\sqrt[3]{\varepsilon})^i, \quad (5)$$

which satisfy Eq. (3) or Eq. (4). Unfortunately, the Jacobian of the two scalar equations in (3) with respect to $t_{11}$ and $T$ is zero at $\varepsilon = t_{11} = T = 0$. A theorem due to Artin (see Section 5 for exact formulation) states that the series (5) have positive radii of convergence if their coefficients can be determined by substituting (5) in (4), after which a series in $\sqrt[3]{\varepsilon}$ results of which all coefficients must be zero. The coefficients of the first terms give

$$b_1(2x_1 - \beta_1) = 0, \quad (6)$$
$$x_0 + b_1(2x_2 - \beta_2) - (Ab_1/2!)(2x_1^2 - \beta_1^2) = 0, \quad (7)$$
$$b_1(2x_3 - \beta_3) - (Ab_1/2!)(4x_1x_2 - 2\beta_1\beta_2) + (A^2b_1/3!)(2x_1^3 - \beta_1^3) = 0, \quad (8)$$
$$b_1(2x_4 - \beta_4) - (Ab_1/2!)(4x_1x_3 - 2\beta_1\beta_3 + 2x_2^2 - \beta_2^2)$$
$$+ (A^2b_1/3!)(6x_1^2x_2 - 3\beta_1^2\beta_2) - (A^3b_1/4!)(2x_1^4 - \beta_1^4) = 0. \quad (9)$$
From these equations, \( \alpha_i \) and \( \beta_i \) can be determined, provided that

\[
\det[b_1, Ab] \neq 0, \quad (10)
\]

\[
\det[b_1, -x_0]/\det[b_1, -Ab] > 0. \quad (11)
\]

Considered in more detail, it follows from (6) that \( 2\alpha_1 = \beta_1 \). The quantity \( 2\alpha_1^2 - \beta_1^2 \) can be solved from (7), and then \( \alpha_1 \) and \( \beta_1 \) can be calculated and then substituted in (8) and (9). Now, the solutions of the two linear systems of equations (7) and (8) give two different linear combinations of \( \alpha_2 \) and \( \beta_2 \). Hence, \( \alpha_2 \) and \( \beta_2 \) can be calculated and then their values are substituted in (9), etc.

If in (11) the \( < \) sign would have been valid, then the t.o.c. would start with a minus sign. Only \( x_0 \) for which \( \det[b_1, x] = 0 \) are excluded from consideration.

For \( \epsilon \) sufficiently small, we have found a bang–bang control of which the optimality still must be proved. We use the following sufficiency condition (see also Section 8): if a control steers the system from its initial to its final position and it can be written as \( \text{sign}(\eta, \exp(-tA)b_1) \) for some nonzero 2-vector \( \eta \), then it is time optimal. Here, the notation \( (c, d) \) denotes the inner product of the vectors \( c \) and \( d \). It remains to find a vector \( \eta \) in such a way that \( \text{sign}(\eta, \exp(-tA)b_1) \) coincides with the control found. For \( \epsilon \) and hence \( T \) very small, we write

\[
\text{sign}(\eta, \exp(-tA)b_1) \approx \text{sign}(\eta, b_1 - tAb_1), \quad 0 \leq t \leq T. \quad (12)
\]

By choosing \( \eta \) in such a way that \( (\eta, b_1) \) is just positive and \( (\eta, Ab_1) \) is positive, then \( (\eta, b_1 - tAb_1) \) is a decreasing linear function of time; and, by adjusting \( \eta \), we can choose the root of this function everywhere on \( 0 \leq t \leq T \). This proves the time optimality of the bang–bang control found.

**Remark 2.1.** Note that, as a first approximation of switching and final time, the ratio of \( t_{11} \) and \( T \) equals 0.5, which is independent of \( x_0 \)!

This can be explained as follows. The optimal trajectory along which the system is transferred from \( \epsilon x_0 \) to the origin does not remain within an \( \epsilon \)-environment of the origin. At first, the system moves away, until a distance of order \( \sqrt{\epsilon} \) is reached, which is very long with respect to the deviation \( \epsilon x_0 \) to be compensated, and then returns to the origin. The system moves along both parts of the trajectory with approximately constant velocity in the \((x_1, x_2)\)-plane and, hence, the position of \( \epsilon x_0 \) does not influence the ratio of \( t_{11} \) and \( T \) very markedly.

**Remark 2.2.** The values of the radii of convergence of the series (5) depend, among other things, on the direction of \( x_0 \). Let the curve \( l_1 \),
see Fig. 2, denote the switching curve in a neighborhood $\Omega$ of the origin. Apart from the points which lie on it, $l_1$ divides $\Omega$ into two open parts (Ref. 3); if $\varepsilon x_0$ lies in one part, the t.o.c. will start with a plus sign; and, if $\varepsilon x_0$ lies in the other part, the t.o.c. will start with a minus sign. If $\varepsilon x_0$ lies on $l_1$, no switch occurs, and again two possibilities exist. The origin divides $l_1$ into two parts; in one part, the plus sign occurs; and, in the other part, the minus sign occurs.

Optimal trajectories always approach the origin along a part of curve $l_1$ (Ref. 3). The tangent to the curve $l_1$ at the origin has the direction of the vector $b_1$. Suppose that $x_0$ lies on the half-line $m$ (see Fig. 2). This line has the origin and point $P$ in common with $l_1$. An elementary calculation shows that the derivative of $T$ with respect to the parameter $\varepsilon$ does not exist at point $P$. If point $P$ corresponds to $\varepsilon = \varepsilon^*$, the radius of convergence of the power expansion of $T$ in Eq. (5) is equal to or smaller than $\varepsilon^*$.

If the half-line $m$ approaches the tangent along the smallest angle, then the radii of convergence of the series (5) tend to zero. So, in this context, there is no sense in choosing $x_0$ on the tangent. It is these $x_0$ which have been excluded! In this way, it is clear that, if $\varepsilon x_0$ lies in the shaded region of Fig. 2, the main theorem may not be applied.

3. $L^*$-Systems

The integer $k$ will be defined as

$$ k = n - 1 - (p - 1)r, $$

(13)
where $p$ has already been defined in Section 1. The integers $p$ and $k$ will appear to be important in the rest of this paper. It is easily seen that $0 \leq k \leq r - 1$.

**Definition 3.1.** Systems such that rank $[B] = r$ are called $L^*$-systems if the following square matrices are nonsingular:

$$L_{i_1, \ldots, i_p, i_{k+1}}^* = [B, AB, \ldots, A^{p-2}B, A^{p-1}b_{i_1}, \ldots, A^{p-1}b_{i_k}, A^q b_{i_{k+1}}],$$

for all numbers $\{i_1, \ldots, i_k, i_{k+1}\}$ in such a way that (i) the numbers $\{i_1, \ldots, i_k\}$ are different numbers and belong to $\{1, \ldots, r\}$; (ii) the number $i_{k+1}$ is equal to one of the numbers $\{1, \ldots, r\}$; if $i_{k+1}$ is equal to one of the numbers $\{i_1, \ldots, i_k\}$, then $q$ is defined as $q = p$ and otherwise $q = p - 1$.

If $k = 0$, the columns $Ab_{i_1}, \ldots, Ab_{i_k}$ do not appear in (14); in this case, $q = p - 1$.

If $p = 1$, the $L^*$-matrix in (14) is defined as

$$L_{i_1, \ldots, i_p}^* = [b_{i_1}, \ldots, b_{i_k}, A^q b_{i_{k+1}}].$$

Obviously, all $L^*$-systems are controllable and have the smallest possible controllability index (Ref. 4), i.e., the index equals $p$. For $r = 1$, the class of $L^*$-systems coincides with the class of controllable systems.

The $L^*$-systems have been introduced for technical reasons: the author could prove the results of this paper for $L^*$-systems only. It is not clear that the class of $L^*$-systems has some physical interpretation. See also Remark 4.2.

### 4. Formulation of the Main Theorem

The solution of the differential equation (1) can be written as

$$x(t) = \exp(tA) \left\{ \varepsilon x_0 + \int_0^t \exp(-sA) Bu(s) \, ds \right\}.$$

(15)

For the time being, we suppose that the t.o.c, which transfers the system from $x_0$ to the origin exists, is unique, and is bang–bang. In the proof of the main theorem, it appears that only systems and initial conditions are considered for which this is true. If the t.o.c. is substituted in (15), we obtain

$$\varepsilon x_0 + \sum_{i=1}^r \left\{ \pm \left( \int_0^{t_{i+1}} - \int_{t_{i-1}} x_{i-1} + \cdots + (-1)^{n_i} \int_{t_{i_{n_i}}}^{T} \right) \exp(-sA) b_i \, ds \right\} = 0,$$

(16)
where $t_{ij}$ is the time at which the $j$th switch of the $i$th control component $u_i$ occurs and $n_i$ is the number of switches of the $i$th control component. It will appear that for many systems, many values of $x_0$, and $\epsilon$ sufficiently small, the total number of switches equals $n - 1$. Let us restrict ourselves to this case for a moment. Then, the vector equation (16) consists of $n$ equations with $n$ unknowns, viz., the switching times and the final time. Also, the unknown signs with which the $u_i(t)$ start (the starting signs) have to be determined.

**Remark 4.1.** One might think that these $n$ equations determine the signs and the unknowns uniquely; in general, this is not true. Even if we suppose the natural restriction

$$0 < t_{i1} < t_{i2} < \cdots < t_{in_i} < T, \quad i = 1, \ldots, r,$$

more solutions may exist. However, we supposed that the t.o.c. was unique and, hence, it has to be extracted from all the solutions obtained from Eq. (16). Of course, the solution with the smallest value of $T$ is the right one.

In the main theorem, an explicit criterion will be given about the numbers of switches and about the starting signs of the different control components. To this end, we need the concept of characteristic matrices. For $L^*$-systems the (square) characteristic matrix $M_{j_1, \ldots, j_r}$ (which is called an $M$-matrix) is defined as

$$M_{j_1, \ldots, j_r} = [J, (-A)^{p-1} b_{j_1}, \ldots, (-A)^{p-1} b_{j_k}, (-A)^{p-1} (\pm b_{j_{k+1}} \pm \cdots \pm b_{j_r})],$$

where

$$J = [B, -AB, \ldots, (-A)^{p-2} B],$$

and where the numbers $j_1, \ldots, j_r$ are an arbitrary permutation of the numbers $\{1, 2, \ldots, r\}$ with the two separate restrictions

$$j_1 < j_2 < \cdots < j_k, \quad j_{k+1} < j_{k+2} < \cdots < j_r.$$

If $p = 1$, the matrix $J$ is not defined and does not appear in the $M$-matrix (17). We will now discuss the choice of the $\pm$ signs in the last column of the $M$-matrix (17). To this end, we consider the vector $\varphi_0$, $\|\varphi_0\| = 1$, defined to be perpendicular to the first $n - 1$ columns of $M_{j_1, \ldots, j_r}$, which are linearly independent. Of the two remaining possibilities for $\varphi_0$ (opposite directions), one is chosen and fixed. Later on, a criterion will be formulated which decides on the two possibilities, but at this moment this is irrelevant. The signs of $(-A)^{p-1} b_{j_i}$, $i = k + 1, \ldots, r$, in the last
column of this characteristic matrix, are now determined by sign 
\((\psi_0, A^{p-1}b_j)\). Because we restricted ourselves to \(L^*\)-systems, these sign 
relations are well defined. Another interpretation of the choice of the 
\(\pm\) signs can be given. The \(\pm\) signs are chosen in such a way that the 
inner product of \(\psi_0\) and \(A^{p-1}(\pm b_j)\) has its maximal value. 
From this interpretation, it follows that the characteristic matrix \(M_{j_1...j_k}\) 
is nonsingular.

Obviously, \(\binom{n}{k}\) different \(M\)-matrices exist for a system, different in 
the sense of different sequences of indices. When there is no risk of 
confusion, the indices of the \(M\)-matrices will be omitted (for \(k = 0\), the 
\(M\)-matrix does not have indices at all).

**Definition 4.1.** The notation \(M[j; x_0]\) denotes a matrix equal to 
the matrix \(M\), except in the \(j\)th column, which is replaced by the column 
vector \(x_0\).

In the proof of the main theorem, it appears that, to almost all unit 
vectors \(x_0\), a unique \(M\)-matrix corresponds in such a way that

\[
| \det[M((p - 1)r + j; x_0)] | < | \det[M(n; x_0)] |, \quad j = 1, \ldots, k, \quad (19-1)
\]

if \(k \geq 1\), and

\[
0 < | \det[M(n; x_0)] | \quad (19-2)
\]

if \(k = 0\). Those \(x_0\) for which such an \(M\)-matrix exists are called admissible. It will be shown that almost all \(x_0\) are admissible. Those \(x_0\) for which 
no \(M\)-matrix exists to satisfy (19) are situated on one or more of a finite 
number of hyperplanes in \(\mathbb{R}^n\). It will appear that these hyperplanes are 
the tangent planes at the switching hypersurfaces at the origin. In the 
main theorem, only admissible \(x_0\) will be considered. For these \(x_0\), the 
unit vector \(\psi_0\) is now uniquely determined by

\[
(\psi_0, x_0) < 0. \quad (20)
\]

Because the right-hand side (r.h.s.) of (19) is greater than zero, it follows 
that \((\psi_0, x_0) \neq 0\) and, hence, \(\psi_0\) is uniquely determined by (20).

**Theorem 4.1. Main Theorem.** If (i) the system (1) is an \(L^*\)-
system, (ii) the unit vector \(x_0\) is admissible, (iii) \(0 < \epsilon < \epsilon_1\), with \(\epsilon_1\) 
sufficiently small (\(\epsilon_1\) may depend on \(x_0\)), and if \(M^*\) is the characteristic 
matrix which satisfies (19) for the vector \(x_0\) mentioned in (ii), then the 
following results hold:

(a) The t.o.c. which steers the system from \(\epsilon x_0\) to the origin is
unique, and the indices which belong to $M^*$ are exactly the indices of the control components which have $p$ switches. The remaining control components have $p - 1$ switches. So together there are

$$pk + (p - 1)(r - k) = n - 1$$

switches.

(b) The starting sign of the control component $u_j(t)$, which has $n_j ( = p$ or $p - 1$) switches, is

$$\text{sign}(\varphi_0, A^{n_j} b_j), \quad j = 1, \ldots, r,$$

where $\varphi_0$ corresponds to the matrix $M^*$.

(c) If $0 \leq \epsilon < \epsilon_1$, where $\epsilon_1$ has been defined in (iii), then

$$t_{ij} = f_{ij}(\sqrt[+]{\epsilon}), \quad T = f(\sqrt[+]\epsilon),$$

where $\sqrt[+]\epsilon$ is to be taken real and nonnegative, and where $f_{ij}(\cdot)$ and $f(\cdot)$ are analytic functions of their argument in a neighborhood of the origin, and where $f_{ij}(0) = f(0) = 0$. In the power series of $f_{ij}$ and $f$, only real coefficients appear, which depend on $A$, $B$, and $x_0$.

Remark 4.2. The arguments of the sign relations in (21) are all $0$, because the systems are $L^*$-systems. The main theorem has been formulated under sufficient conditions, but simple examples indicate that these conditions can be weakened. It is to be expected that the statement of the main theorem remains more or less valid for controllable systems with the smallest possible controllability index $\nu_c$, i.e., $\nu_c = p$. The uniqueness of the optimal control cannot be guaranteed in general. Furthermore, it is to be expected that the main theorem can be generalized to controllable systems with $\nu_c > p$. The switches and final final time will be analytic functions of $\sqrt[+]\epsilon$, where $\epsilon$ denotes the distance from the initial perturbation to the equilibrium position. The total number of switches will increase with $\nu_c$.

Remark 4.3. The main theorem states that the optimal control is unique; hence, the point $-\epsilon x_0$ is an exposed point of the reachable set $R(T)$ (Ref. 2, p. 63). Only for normal systems (Ref. 2, p. 65) are all boundary points of the reachable set exposed. $L^*$-systems are not necessarily normal. Hence, for an $L^*$-system, which is not normal, points on the boundary of the reachable set exist which are not exposed and for which the t.o.c. is not unique. However, these points just belong to the $x_0$ excluded from consideration.
Because the proof of the main theorem is rather long, it will be given rather concisely. Anyhow, all the essential features are given. The extensive proof has been given in Ref. 5. The proof is spread over Sections 5, 6, 7, 8, and 9. In Section 5, some preliminary results are mentioned, including Artin’s theorem. In Sections 6 and 7, it is shown that formal power series (Ref. 6, f.p.s.) of the switching times and the final time in the variable \( \sqrt{\epsilon} \) exist, which satisfy (16).

In Section 6, the coefficients of the first terms of the f.p.s. are considered. Section 7 deals with the coefficients of the higher-order terms. The theorem of Artin is used to prove that the f.p.s. have positive radii of convergence [unfortunately, the Jacobian of the functions (16) with respect to \( t_{ij} \) and \( T \) equals zero at \( t_{ij} = T = \epsilon = 0 \), except for the case \( r = n \)]. In this way, a bang–bang control has been found which steers the system from \( \epsilon x_0 \) to the origin. In Section 8, it will be proved that this control is time optimal. In Section 9, it will be shown that, for almost all \( x_0 \), restricted to \( \| x_0 \| = 1 \), an \( M \)-matrix can be found which satisfies (19).

5. Some Preliminary Results

**Theorem 5.1.** We are given the functions \( f_j(z_1, \ldots, z_n; w) \), \( j = 1, \ldots, n \), which are analytic in a neighborhood of \( z_1 = \cdots = z_n = w = 0 \). Suppose that f.p.s. exist,

\[
\zeta_i = \sum_{j=1}^{\infty} b_{ij} w^j, \quad i = 1, \ldots, n, \tag{22}
\]

which satisfy

\[
f_j(\zeta_1, \ldots, \zeta_n; w) = 0, \quad j = 1, \ldots, n,
\]

where the f.p.s. \( f_j(\zeta_1, \ldots, \zeta_n; w) \) have been obtained by formal substitution of (22) into the power series expansion of \( f_j \) about the point \( z_1 = \cdots = z_n = w = 0 \). If an integer \( N \) exists in such a way that, for \( j \geq N \), the coefficients \( b_{ij} \), \( i = 1, \ldots, n \), are uniquely determined by

\[
f_i(\zeta_1, \ldots, \zeta_n; w) = 0, \quad i = 1, \ldots, n,
\]

then the series (22) have positive radii of convergence.

The proof will be omitted because Theorem 5.1 is a special case of Theorem 1 in Ref. 1.
Consider the functions
\[ g_j(\alpha_1, \ldots, \alpha_{n-1}; a) = (-1)^{n-1} - a \cdot \delta_{j,n-1} + 2 \sum_{i=1}^{n-1} \{(-1)^{i-1} \alpha^i\}, \quad j = 1, \ldots, n, \]
where \( \delta_{j,n-1} \) denotes the Kronecker symbol and \( a \) is a real number. Only real quantities will be considered with respect to these equations.

**Theorem 5.2.** The system of equations
\[ g_j(\alpha_1, \ldots, \alpha_{n-1}; 0) = 0, \quad j = 1, \ldots, n - 1, \]
has a solution of the form
\[ \alpha_i = \sin^2(i\pi/2n), \quad i = 1, \ldots, n - 1; \quad (23) \]
and, if this solution is substituted in \( g_n(\alpha_1, \ldots, \alpha_{n-1}; 0) \), the result is \( n/(-4)^{n-1} \).

The proof follows by nontrivial substitution. \( \square \)

**Theorem 5.3.** If the restriction \( 0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{n-1} \leq 1 \) is added to Theorem 5.2, then the solution (23) is unique.

**Proof.** Consider system (1) with
\[ A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \]
This system is normal (Ref. 2) and, hence, the t.o.c. which steers the point \((e, 0, \ldots, 0)\) to the origin is unique (Ref. 2, p. 69). It is easily verified that the switching times \( t_{\alpha} \) and the final time \( T \) of this control satisfy \( g_j(\alpha_1, \ldots, \alpha_{n-1}; 0) = 0, \quad j = 1, \ldots, n - 1, \) with \( \alpha_i = t_{\alpha}/T \). Because the t.o.c. is unique, the solution of \( g_j(\alpha_1, \ldots, \alpha_{n-1}; 0) = 0 \) is unique too within the restriction \( 0 \leq \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_{n-1} \leq 1 \). \( \square \)

With the aid of the well-known implicit function theorem and Theorems 5.2 and 5.3, the following theorem can be proved.

**Theorem 5.4.** For \( |a| \leq (n - 1)/4^{n-2} \), the system of equations
\[ g_j(\alpha_1, \ldots, \alpha_{n-1}; a) = 0, \quad j = 1, \ldots, n - 1, \]
has a unique solution $\alpha_i(a)$, subject to the restriction
\begin{equation}
0 \leq \alpha_1(a) < \alpha_2(a) < \cdots < \alpha_{n-1}(a) \leq 1.
\end{equation}
If $|a| < (n-1)/4$, the equality signs in (24) may be removed. The solutions $\alpha_i(a)$ are monotonic on the interval $|a| \leq (n-1)/4$.

6. First Term in the Expansions of the Switching Times and Final Time

In this section and the next, we will show that f.p.s. exist of the switching times and the final time which satisfy Eq. (16). In this section, the first term of these f.p.s. will be considered.

We assume that the conditions of the main theorem have been satisfied and that an $M$-matrix exists in such a way that Ineqs. (19) are valid. In Section 9, the latter assumption will be proved to hold for almost all $x_0$. In Section 4, it was stated that the control components corresponding to the indices of this $M$-matrix have $p$ switches and that the remaining control components have $p-1$ switches. We assume this to be true for the moment (it will be proved in Section 8) and will construct f.p.s. which satisfy these conditions.

For convenience of notation, we assume that the column vectors $b_i$ are ordered in such a way that the control components corresponding to $b_1,\ldots, b_k$ have $p$ switches and that the components corresponding to $b_{k+1},\ldots, b_r$ have $p-1$ switches. In addition, we assume that the starting signs of all control components are positive, i.e., $\text{sign}(\vec{v}_0,A^n b_i) = +1$, $i = 1,\ldots, r$. This can always be realized by multiplying a column vector $b_i$ by the factor $-1$, if necessary. The relations (19) and (21) are not disturbed by these multiplications. After these simplifications, the nonsingular $M$-matrix which satisfies Ineqs. (19) is equal to
\begin{equation}
M = \left[ J, (-A)^{p-1} b_1, \ldots, (-A)^{p-1} b_k, \sum_{v=k+1}^r (-A)^{p-1} b_v \right].
\end{equation}
The assumptions about the switch distribution and the starting signs will hold for Sections 6, 7, 8, and 9. Also for reasons of notation, $\epsilon (\geq 0)$ is replaced by $\epsilon_0^p$, where $\epsilon_0$ is a nonnegative quantity.

Starting from these assumptions, we can write Eq. (14), after having developed $\exp(-sA)$ in its Taylor expansion about $s = 0$ and after having carried out the integrations, as
\begin{equation}
\sum_{i=1}^r \left[ \sum_{n=1}^\infty \left( -1 \right)^{n-1} \left( A^{n-1} b_i / \mu ! \right) \left( 2 \sum_{j=1}^n \left[ (-1)^{j-1} T_0^j \right] + (-1)^n T^n \right) \right] + \epsilon_0^p \cdot x_0 = 0,
\end{equation}
where \( n_v = p \) if \( v \leq k \) and \( n_v = p - 1 \) if \( v > k \). The f.p.s.
\[
  t_{\nu \lambda} = \sum_{m=1}^{\infty} \alpha_{\nu \lambda m} e_0^m, \quad T = \sum_{m=1}^{\infty} \beta_m e_0^m
\]  
(27)

(the coefficients \( \alpha_{\nu \lambda m}, \beta_m \), are still unknown) are substituted in Eq. (23), after which a f.p.s. in \( e_0 \) results of which the coefficients must be zero. These coefficients are equations in \( \alpha_{\nu \lambda m} \) and \( \beta_m \) which can be solved now.

After some elementary, tedious substitutions, it follows that \( \alpha_{\nu \lambda 1} \) and \( \beta_1 \) satisfy
\[
  \beta_1^p = (p - 1)! \cdot (4)^{p-1} \cdot |\det[M(n; x_0)]/\det[M]|,
\]
and
\[
  (-1)^{p-1} + 2 \sum_{\lambda=1}^{p-1} \{( -1)^{\lambda-1} (\alpha_{\nu \lambda 1}/\beta_1)^{\lambda} \} = 0,
\]
\[
  j = 1, \ldots, p - 1, \quad \nu = k + 1, \ldots, r,
\]  
(28)

\[
  (-1)^{p} - \delta_{i,v}[p/(4)^{p-1}] \chi_v + 2 \sum_{\lambda=1}^{p} \{( -1)^{\lambda-1} (\alpha_{\nu \lambda 1}/\beta_1)^{\lambda} \} = 0,
\]
\[
  j = 1, \ldots, p, \quad \nu = 1, \ldots, k,
\]  
(29)

where
\[
  \chi_v = \det[M((p - 1)r + v; x_0)]/\det[M(n; x_0)].
\]  
(30)

Systems (28) and (29) have been considered in Theorem 4.5, which states that these systems have solutions which satisfy
\[
  0 < \alpha_{\nu 11}/\beta_1 < \alpha_{\nu 21}/\beta_1 < \cdots < \alpha_{\nu p1}/\beta_1 < 1, \quad \nu = 1, \ldots, r,
\]
provided that
\[
  |\chi_v| < 1, \quad \nu = 1, \ldots, k.
\]  
(31)

In Section 9, it will be proved that, for almost all \( x_0 \), an \( M \)-matrix exists to satisfy (31). Now, all coefficients \( \alpha_{\nu \lambda 1} \) and \( \beta_1 \) have been determined.

Note that, if \( k = 0 \), no \( \nu \) exists which satisfies \( \nu = 1, \ldots, k \); in that case, the statement concerning these \( \nu \) is a dummy statement. This can happen repeatedly in the next sections.

### 7. Higher-Order Terms

In this section, it will shown that the coefficients \( \alpha_{\nu \lambda m}, \beta_m \) for \( m \geq 2 \) are uniquely determined by recurrence relations. The proof is given by
induction. It is assumed that we already know the coefficients \( \alpha_{v+m}, \beta_{v} \), \( 1 \leq m \leq q - 1 \), for some \( q \geq 2 \); and, starting from there, it can be shown that the coefficients \( \alpha_{v}, \beta_{v} \) can be determined and expressed in \( \alpha_{v+m}, \beta_{v}, 1 \leq m \leq q - 1 \).

It is only a matter of some laborious book-keeping to show that

\[
Mv_{q} = \xi_{q},
\]

\[
(-1)^{h} h^{\alpha_{v}^{h-1}} \beta_{v} + 2 \sum_{i=1}^{q} \{(-1)^{h-1} h^{\alpha_{v}^{h-1}} \alpha_{v}\} = \xi_{q, v, h},
\]

\( h = 1, \ldots, p, \quad v = 1, \ldots, k \) \hspace{1cm} (33)

\[
(-1)^{v-1} h^{\alpha_{v}^{v-1}} \beta_{v} + 2 \sum_{i=1}^{v-1} \{(-1)^{h-1} h^{\alpha_{v}^{v-1}} \alpha_{v}\} = \xi_{v, q, h},
\]

\( h = 1, \ldots, p, \quad v = k + 1, \ldots, r \) \hspace{1cm} (34)

Here, \( v_{q} \) and \( \xi_{q} \) are \( n \)-vectors and \( \xi_{v, q, h}^{*} \) are real numbers. The vector \( \xi_{q} \) and the numbers \( \xi_{v, q, h}^{*} \) are functions of \( \alpha_{v+m}, \beta_{v} \), \( 1 \leq m \leq q - 1 \), and hence are known. The matrix \( M \) is (32) equals the \( M \)-matrix in (25) and, hence, is nonsingular. The unknown vector \( v_{q} \) can be solved, and it appears that the last component of \( v_{q} \) equals \( \beta_{k} \). In this way, \( \beta_{k} \) is uniquely determined and it is assumed to be known now.

The unknown coefficients \( \alpha_{v}, \beta_{v} \) appear linearly in Eqs. (33) and (34). It is easily checked that the system matrices of these linear equations show a Van der Monde character and are nonsingular on account of \( \alpha_{v}^{1} < \alpha_{v}^{2} < \cdots < \alpha_{v}^{n+1} < \beta_{1} \). Hence, \( \alpha_{v}, \beta_{v} \) are uniquely determined by (33) and (34).

8. Solution Found is Optimal

After substitution of the series (27) in (26) the coefficients of these series have been determined in Sections 6 and 7, and now Theorem 4.1 can be applied. The conditions of this theorem (take \( N = 1 \)) have been satisfied. Hence, the series (27) are convergent for \( \epsilon_{0} \) sufficiently small. For these \( \epsilon_{0} \), a bang-bang control has been found which transfers \( \epsilon x_{0} = \epsilon_{0} x_{0} \) to the origin. In general, it is not necessary for this control to be time optimal. However, if a unit \( n \)-vector \( \eta \) can be found in such a way that the bang-bang control is identical with

\[
u_{i}(t) = \text{sign}(\eta, \exp(-tA) b_{i}), \quad i = 1, \ldots, r,
\]
then the solution really represents the t.o.c. (Ref. 2, Theorem 13.2). In this section, we will show that such a vector $\eta$ exists.

**Theorem 8.1.** For $\epsilon$ sufficiently small, the vectors $\exp(-t_{ij}A) b_i$, $i = 1, \ldots, r$, $j = 1, \ldots, n_i$, are linearly independent.

**Proof.** Define the $n \times n$ matrix $Z$ as

$$Z = [\exp(-t_{11}A) b_1, \ldots, \exp(-t_{rn_A}A) b_r, \tilde{\psi}_0],$$

(36)

where $\tilde{\psi}_0$ corresponds to the $M$-matrix in (25). It appears that

$$\det[Z] = \prod_{k=1}^{r} \left( \prod_{j=1}^{n_r} \left( (\alpha_{k11} - \alpha_{kj}) \epsilon_0 + O(\epsilon_0^2) \right) \right),$$

(37)

$$\det[b_1 + O(\epsilon_0), (-A)b_1, (-A)^{n_r-1}b_r/(n_r-1)! + O(\epsilon_0), \tilde{\psi}_0].$$

(38)

For $\epsilon_0$ sufficiently small and positive, the r.h.s. is nonzero and, hence, $Z$ is nonsingular, which proves the theorem. \[\square\]

From this theorem, it follows that a unit $n$-vector $\psi_0(\epsilon)$ exists, which is, apart from a factor $\pm 1$, uniquely determined by

$$(\psi_0(\epsilon), \exp(t_{ij}(A)b_i) = 0, \quad i = 1, \ldots, r, \quad j = 1, \ldots, n_i.$$ (39)

**Theorem 8.2.** The vector function $\psi_0(\epsilon)$ is continuous in $\epsilon$ and

$$\lim_{\epsilon \to 0} \psi_0(\epsilon) = \pm \tilde{\psi}_0,$$

(40)

where $\tilde{\psi}_0$ has been defined in Section 4; $\tilde{\psi}_0$ corresponds to the $M$-matrix (25).

**Proof.** For small $\epsilon$, Eq. (39) can be written as

$$(\psi_0(\epsilon))' [b_1 + O(\epsilon), \ldots, (-A)^{n_r-1}b_r/(n_r-1)! + O(\epsilon)] = 0.$$ (41)

Eq. (40) is a direct result of Eq. (41). \[\square\]

We will now choose that direction of $\psi_0(\epsilon)$ for which the $+$ sign is valid in (40). On account of (20), we have $(\psi_0(\epsilon), x_0) < 0$ for $\epsilon$ sufficiently small. This coincides with the theory treated in Ref. 2, p. 51.

**Theorem 8.3.** For $\epsilon$ sufficiently small, the functions $h_i(\epsilon, t)$, defined by

$$h_i(\epsilon, t) = (\psi_0(\epsilon), \exp(-tA)b_i), \quad i = 1, \ldots, r,$$
have precisely $n_i$ zeros, respectively, on the interval $0 < t < T(\epsilon)$. These zeros are $t_{i1}, ..., t_{in_i}$; and, at these points, the function $h_i(\epsilon, t)$ changes sign. Moreover, with regard to the starting signs, the following formulas are valid:

$$\text{sign}(f_i(\epsilon, 0)) = \text{sign}(\psi_0(\epsilon), b_i) = \text{sign}(\psi_0, A^\nu b_i), \quad i = 1, ..., r.$$  

**Proof.** $h_i(\epsilon, t)$, as a function of $t$, behaves essentially as a polynomial of degree $n_i$ on the interval $0 \leq t \leq T(\epsilon)$, for fixed $\epsilon$. This polynomial has $n_i$ zeros and it can be proved that the $n_i$th derivative of $h_i(\epsilon, t)$ with respect to $t$ has a constant sign on $0 \leq t \leq T(\epsilon)$.

Now, set $\eta = \psi_0(\epsilon)$; then, it follows that the bang–bang control found is identical with the r.h.s. of (35), from which it follows that this control is time optimal.

**Remark 8.1.** Because $h_i(\epsilon, t)$ has a finite number of zeros on $[0, T(\epsilon)]$, the $i$th control component is uniquely determined almost everywhere by $\text{sign}(h_i(\epsilon, t))$ for $i = 1, ..., r$, and hence the t.o.c. is unique (Ref. 2, pp. 61 and 63).

9. Main Theorem is Applicable to Almost All $x_0$

In this section, we will prove that, for almost all unit vectors $x_0$, a matrix $M$ can be found in such a way that

$$|\det[M{(p - 1)\nu + \nu; x_0}]/\det[M{\nu; x_0}]| = |x_\nu| < 1, \quad \nu = 1, ..., k.$$  

(42)

According to (19), the denominator of the left-hand side of (42) must be $\neq 0$.

**Theorem 9.1.** For each unit vector $x_0$, at most one $M$-matrix exists, which satisfies (42).

**Proof.** The existence of two different $M$-matrices, both satisfying (42), would imply that two different t.o.c. of the control problem exist, which is contradicted by the uniqueness of the t.o.c. according to Remark 8.1.
to the unit ball \( \| x \| = 1 \) of points \( x_0 \) can be found in such a way that Ineqs. (42) are valid.

Proof. If an \( M \)-matrix is given, choose \( x_0 \), to be denoted by \( x_0 \), equal to the (normed) last column of this \( M \)-matrix, or choose \( x_0 = -x_0 \). In these cases, it is clear that (42) is satisfied. The determinants in (42) are continuous in \( x_0 \); hence, Ineqs. (42) are not only satisfied for the \( \pm x_0 \) mentioned, but for all points \( x_0 \) is a neighborhood of these \( \pm x_0 \).

Consider the unit ball \( \| x \| = 1 \) in \( \mathbb{R}^n \) and on this ball the set of those \( x_0 \) for which (42) is valid and, in addition, for which \( \det[M(n; x_0)] \) has a constant sign. Suppose, for the time being, that this sign is positive. This set will be denoted by \( X^0_{M^+} \). From Theorem 9.2, it follows directly that the set of these \( x_0 \) is open and nonempty. With respect to the unit ball, the set \( X^0_{M^+} \) is even convex, which is easily shown.

The set \( X^0_{M^+} \) will now be called a domain. In the same way, the set \( X^0_{M^-} \), i.e., the set of all \( x_0 \) which satisfy (42) and for which \( \det[M(n; x_0)] \) has a negative sign, is a domain. Hence, those points \( x_0 \) which satisfy (42) for a certain \( M \)-matrix can be divided into two domains. The point \( x_0 = x_0 \) and its neighborhood belong to one domain, and the point \( x_0 = -x_0 \) with its neighborhood belong to the other. The boundary of such a domain on the unit ball consists of those points \( x_0 \) for which in (42) the equality symbol appears for some of the \( v \) instead of the \( < \) symbol. In other words, the boundary consists of the points of intersection of the unit ball with parts of linear subspaces of dimension \( n - 1 \). Such an intersection will be called a boundary part. The number of these linear subspaces is less than or equal to \( 2k \). The denominator of \( x^r, \det[M(n; x_0)] \), has a constant sign, and the numerator

\[
\det[M((p - 1)r + v; x_0)]
\]

may be positive or negative. It is of course not necessary that all these \( 2k \) subspaces should really appear as a boundary part of one domain.

Because only a finite number of \( M \)-matrices exist, only a finite number of domains exists. Let such a domain be denoted by \( G_i \). It can be proved that, if we cross an arbitrary boundary part from inside \( G_i \) to outside \( G_i \), then we enter another domain. Or, to put it differently, an arbitrary boundary part of \( G_i \) is also boundary part of some other domain \( G_j \). Because \( G_i \) and \( G_j \) are disjunct, they are neighbors. This proves that the whole unit ball (except the boundary parts) is covered with domains. The domains induce a honeycomb-like structure on the unit ball. Because only a finite number of boundary parts exists, it follows that almost all \( x_0 \) belong to a domain.
10. Relation of the Results Obtained to Canonical Forms

In this section, only \((n, 1)\) systems are considered. A system related to system (1) will be given, with a related initial condition, in such a way that the first terms in the expansions (27) are the exact solution to the time-optimal control problem of this related system and initial condition.

Considering only the first terms in the expansions of the switching times and the final time (for the original problem), we find:

\[
T \approx ((n - 1)! \cdot (-4)^{n-1} \cdot \frac{\det[M(n; -ex_0)]}{\det[M]}^{1/n},
\]

\[
t_{1;i} \approx T \sin^2(i\pi/2n), \quad i = 1, ..., n - 1,
\]

where

\[
M = [b_1, -Ab_1, ..., \pm(-A)^{n-1}b_1].
\]

Note that the ratio of these first-order approximations is independent of the direction vector \(x_0\) (see Remark 2.1).

By means of a proper linear coordinate transformation (Ref. 4) \(x \rightarrow \tilde{x}\), the equations of motion can be written in the canonical form

\[
\frac{d\tilde{x}}{dt} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \vdots \\ x & x & x & \cdots & x \end{bmatrix} \tilde{x} + \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} u, \quad \tilde{x}(0) = \varepsilon \tilde{x}_0 \triangleq \varepsilon \begin{bmatrix} \tilde{x}_{01} \\ \vdots \\ \vdots \\ \tilde{x}_{0n} \end{bmatrix},
\]

where the \(x\) in the system matrix denotes possibly nonzero elements. Now, consider the related system and related initial condition

\[
\frac{d\tilde{x}}{dt} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix} \tilde{x} + \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix} u, \quad \tilde{x}(0) = \varepsilon \begin{bmatrix} \tilde{x}_{01} \\ 0 \\ \vdots \\ \tilde{x}_{0n} \end{bmatrix}.
\]

The t.o.c. corresponding to (46) is exactly known (see the proof of Theorem 5.3); the switching times and the final time are equal to the corresponding expressions in Eq. (43), where the \(\approx\) symbol for this case is replaced by the \(=\) symbol. Hence, one can state that, as far as the first approximations of the switching times and the final time are concerned, the system under consideration may be considered as a system of the form (46), which is called a perfect \(n\)-fold integrator.
11. Numerical Aspects

Since the switching times and final time are analytic functions of $\epsilon_0 = \sqrt{\epsilon}$ for $\epsilon$ sufficiently small, provided that the system and $x_0$ satisfy the conditions of the main theorem, the expansions of these times can be used for numerical purposes.

The known methods, see for instance (Refs. 7–8), are mostly only applicable to systems where $n$ and $r$ are rather small. Even in these cases, one can meet with many difficulties when solving such a time-optimal control problem. The method of expansion of the switching times and final time seems to be applicable also to systems where $n$ (and possibly $r$) are large. However, a limitation is now that the series have to converge, hence that $\epsilon$ must be sufficiently small. If this condition is satisfied, then the computer needs hardly any time to solve the problem.

A difficulty may be that one does not know when $\epsilon$ is sufficiently small; we did not investigate the radii of convergence of the different series. So, if finitely many coefficients of the expansions are calculated, the truncated expansions may seem to converge; one does know this for certain. This uncertainty, however, it not of great importance. If finitely many coefficients are calculated, one can check the exactness of the approximate switching times and final time. The original equations (1) can be integrated if these values are substituted, and one can look to see if the final situation has been achieved within a degree of accuracy desired. That will generally be the practical requirement.

If one needs the switching times and final time very accurate, it is not necessary to calculate more and more coefficients. The Newton–Raphson method can also be applied to Eq. (2), which consists of $n$ scalar equations with $n$ unknowns. In this case, one starts with the method described above; a finite number of coefficients is calculated, and the obtained results will be used as the starting values for the Newton–Raphson procedure.

The author has numerical experience with (4,1) and (4,2) systems. On a Telefunken TR4 computer, the calculation up to seven terms in the expansion could be performed within 1 and 2 seconds, respectively. However, concerning the range of $\epsilon$ for which reasonable answers were obtained, the results are not encouraging.

References