A TWO-POINT METHOD FOR THE NUMERICAL SOLUTION OF SYSTEMS OF SIMULTANEOUS EQUATIONS*

BY

W. M. KINCAID

University of Michigan

1. Introduction. This paper is concerned with the problem, which arises frequently in practice, of finding one or more solutions of a set of n simultaneous equations $\varphi_i(x_1, \cdots, x_n) = 0$, $i = 1, \cdots, n$, in n unknowns $x_1, \cdots, x_n$. For convenience we shall assume, as is usually true in actual cases, that the functions φ_i are analytic in a region surrounding each solution of interest, although the approach we shall use would remain applicable under much milder conditions.

While a wide choice of methods is available when the equations are linear, the same cannot be said in the non-linear case. Simplifying the system by elimination is likely to be laborious, and may be impossible if the equations are not algebraic. If the system is to be tackled in its original form, about the only general procedures are Newton's method and the method of steepest descent. Both of these require the repeated evaluation of all n^2 partial derivatives $\partial\phi_i/\partial x_j$. In addition, Newton's method involves the solution of a set of linear equations at each step. The method of steepest descent does not require this, but generally exhibits only first-order convergence, whereas the convergence of Newton's method is of the second order in general.

These remarks do not mean that existing methods are useless, but they make it clear that there is room for fresh suggestions. In the following section we shall presen- an iterative method, believed to be new, that does not require the calculation of derivt atives or the solution of sets of linear equations, and yet displays second-order convergence. The method may be regarded as a generalization of the classical method of false position for the solution of one equation in one unknown.

While experience with the method has not been extensive, it has been sufficiently encouraging to suggest that the method is worthy of serious examination. It may even be worth considering for the solution of linear systems. In this case the solution is attained (apart from rounding-off errors) after a finite number of steps, and the volume of computation appears comparable in magnitude with that required by existing methods.

2. Description of the method. Since the method of solution to be described is a generalization of the method of false position, as stated earlier, a brief review of the latter is in order.

Let $f(x) = 0$ be an equation in one unknown, and consider first the case where f is linear. If the value of $f(x)$ is computed for two distinct values a and b of x, its value for all x is given by the identity

$$f(x) \equiv \frac{1}{b - a} [(x - a)f(b) + (b - x)f(a)].$$

*Received September 9, 1959. This work was conducted under contract to the United States Army Signal Air Defense Engineering Agency, Contract No. DA-36-039 SC-64627.

Setting the right member equal to zero, one obtains

$$x = \frac{af(b) - bf(a)}{f(b) - f(a)}$$

for the solution of the equation.

If f is non-linear, one can perform the same operation, obtaining the zero of the linear function coinciding with f for $x = a$ and $x = b$. The operation can then be repeated, replacing a or b by the value of x thus obtained. Continuing in the same manner, one obtains a sequence of values of x that will converge to the desired solution if the function is well behaved and the starting points a and b are suitably chosen. The order of convergence of this procedure, which is probably the most efficient form of the classical method of false position, has been shown to be $(1/2)\,[1 + (5)^{1/2}]$ (see [1, 2]) and is thus greater than 1. (By the order of convergence is meant the greatest value of k for which the ratio δ_n/δ_{n-1}^k is bounded, where δ_{n-1} and δ_n are the departures of two successive iterates from the true solution).

Now let f be a real-valued function defined on a region of Euclidean n-space including the points $A(a_1, \cdots, a_n)$ and $B(b_1, \cdots, b_n)$. By analogy with the one-dimensional case, we may determine a new point X having the coordinates

$$X_i = \frac{a_i f(B) - b_i f(A)}{f(B) - f(A)} \qquad (i = 1, 2, \cdots, n) \tag{1}$$

(provided $f(A) \neq f(B)$, as we shall assume in similar cases henceforth). For convenience (1) may be written in the abbreviated form

$$X = AfB. \tag{1a}$$

It is apparent that the point X thus determined lies on the line joining A and B, and it follows from a slight generalization of our earlier statements that $f(X) = 0$ if f is linear; speaking geometrically, X is then the intersection of the line AB and the hyperplane $f = 0$. Moreover, if A and B both satisfy another linear equation $g = 0$, the same is true of X.

Making use of these ideas, one can set up a variety of methods for solving sets of simultaneous linear equations. One such method has been proposed by Purcell [3]. Some of the methods may be applied to sets of non-linear equations as well. It is appropriate to refer to these as *two-point* methods in view of the character of the operation defined by (1) on which the methods are based. The remainder of the present paper is concerned with one method of this type, which possesses some particularly desirable properties.

In the following we shall confine ourselves to the case of two equations

$$\phi_i(X) \equiv \phi_i(x_1, x_2) = 0, \qquad i = 1, 2,$$

in two unknowns, as it is illustrative of the general case. Without loss of generality, we may suppose that $O(0, 0)$ is a solution. We shall assume that the vectors grad $\phi_1(O)$ and grad $\phi_2(O)$ are non-zero and non-collinear, so that the solution is a simple one.

We begin the process of solution by selecting three linear combinations $f(X)$, $g(X)$, $h(X)$ of $\phi_1(X)$ and $\phi_2(X)$ in such a way that $f(X) + g(X) + h(X) \equiv 0$, while any pair of f, g, h are linearly independent. Thus the equations $f(X) = 0$, $g(X) = 0$, $h(X) = 0$ have the common solution O, and the vectors grad $f(O)$, grad $g(O)$ and grad $h(O)$ are non-zero and point in three different directions. The reason for introducing an additional equation will be indicated in a moment.

We select three initial points R_1, S_1, T_1. Making use of the operation defined by (1), we locate successively the points

$$\begin{aligned} S_1' &= R_1 f S_1, \quad & T_1' &= R_1 f T_1, \quad & R_1' &= S_1' g R_1, \\ T_2 &= S_1' g T_1', \quad & R_2 &= T_2 h R_1', \quad & S_2 &= T_2 h S_1'. \end{aligned} \tag{2}$$

The equations (2) represent one iterative cycle. The next cycle proceeds in the same way, with R_2, S_2, T_2 in place of R_1, S_1, T_1.

If the functions f, g, h are linear, the lines $S_1'T_1'$ and R_2T_2 coincide with $f = 0$ and $g = 0$ respectively; thus T_2 (also R_2, S_2) coincides with O. If the functions are non-linear, but not too badly so, one expects these relations to hold approximately, and in this case R_2S_2 should coincide approximately with $h = 0$. Thus if all goes well, successive cycles should yield a sequence of approximately similar triangles converging on O.

The fact that the sides of the triangles approach distinct fixed directions is important. For otherwise one of the triangles might collapse to a line, in which case all succeeding points would lie on the same line, and the solution could not in general be attained. It was with these considerations in mind that the two original equations were replaced by three.

The foregoing discussion is of course quite heuristic. We shall show later that it can be made rigorous if the initial points satisfy suitable conditions.

3. Useful definitions and lemmas. Interpreting a point P in the plane as a vector from O to P, we may define the linear functions

$$u(P) = P \cdot U, \qquad v(P) = P \cdot V, \qquad w(P) = P \cdot W, \tag{3}$$

where

$$U = \text{grad } f(O), \qquad V = \text{grad } g(O), \qquad W = \text{grad } h(O). \tag{4}$$

Thus $u(P)$, $v(P)$, $w(P)$ coincide respectively with the linear terms of the Taylor expansions of f, g, h about O. Hence

$$u(P) + v(P) + w(P) \equiv 0. \tag{5}$$

We see that any pair of the three numbers $u(P), v(P), w(P)$ may be regarded as a linearly transformed set of coordinates of the point P; this interpretation will underlie our subsequent discussion. The 'axes' $u = 0$, $v = 0$, $w = 0$ are the respective tangents to the curves $f = 0$, $g = 0$, $h = 0$ at O.

The non-linear parts of f, g, h are given by the functions

$$F(P) = f(P) - u(P), \qquad G(P) = g(P) - v(P), \qquad H(P) = h(P) - w(P). \tag{6}$$

For any two points P, Q in an appropriate neighborhood of O, we have by the mean value theorem

$$F(P) - F(Q) = (P - Q) \cdot \text{grad } F(\theta P + (1 - \theta)Q) \tag{7}$$

for some θ in $[0, 1]$. Since grad $F(O) = 0$ and F is analytic (in x, y and thus in u, v) near O, there exist positive constants A' and K' such that

$$|\text{ grad } F(P) \,| \leq A' \max (|\, u(P) \,|, |\, v(P) \,|) \tag{8}$$

whenever $\max(|u(P)|, |v(P)|) < K'$. It follows from (7) and (8) that

$$|F(P) - F(Q)| \leq A'\rho(P, Q) \max(|u(P)|, |u(Q)|, |v(P)|, |v(Q)|), \tag{9}$$

(where $\rho(P, Q)$ is the ordinary Euclidean distance between the points P and Q) for $\max(|u(P)|, |u(Q)|, |v(P)|, |v(Q)|) < K'$.

These relations make it convenient to define the norm $\|P - Q\|$ of a vector $P - Q$ by the relation

$$\|P - Q\| = \max(|u(P) - u(Q)|, |v(P) - v(Q)|, |w(P) - w(Q)|). \tag{10}$$

In particular, we write

$$\|P\| = \|P - O\| = \max(|u(P)|, |v(P)|, |w(P)|). \tag{11}$$

Clearly the norm, so defined, has the usual required properties, and the ratio $\|P - Q\|/\rho(P, Q)$ has positive upper and lower bounds. Thus from (9), (10), and (11) we conclude that there exist positive constants A and K such that

$$|F(P) - F(Q)| \leq A\|P - Q\| \max(\|P\|, \|Q\|) \tag{12}$$

for $\max(\|P\|, \|Q\|) < K$. In particular,

$$|F(P)| \leq A\|P\|^2 \quad \text{for} \quad \|P\| < K. \tag{13}$$

Repeating the same arguments, we may suppose A and K so chosen that (12) and (13) hold with G or H in place of F.

We note at this point that (5) and (11) can be combined to yield

$$\|P\| = \min(|u(P)| + |v(P)|, |u(P)| + |w(P)|, |v(P)| + |w(P)|); \tag{14}$$

a similar relation holds for $\|P - Q\|$.

Finally, we introduce

$$r_u(P, Q) = \max\left(\left|\frac{u(P) - u(Q)}{v(P) - v(Q)}\right|, \left|\frac{u(P) - u(Q)}{w(P) - w(Q)}\right|\right), \tag{15}$$

with similar definitions for $r_v(P, Q)$ and $r_w(P, Q)$. We see that $r_u(P, Q)$ may be regarded as measure of the difference in direction between the line PQ and the line $u = 0$; parallelism corresponds to $r_u(P, Q) = 0$.

From (10) and (15), we have

$$|u(P) - u(Q)| \leq r_u(P, Q)\|P - Q\|. \tag{16}$$

On the other hand, if $v(P) \neq v(Q)$, one obtains from (14) and (15)

$$\begin{aligned} \|P - Q\| &\leq |u(P) - u(Q)| + |v(P) - v(Q)| \\ &= |v(P) - v(Q)|\left(\left|\frac{u(P) - u(Q)}{v(P) - v(Q)}\right| + 1\right) \\ &\leq |v(P) - v(Q)|(r_u(P, Q) + 1). \end{aligned} \tag{17}$$

The inequalities (16) and (17) are clearly valid for any permutation of the letters u, v, w. Geometrically, (17) implies that if PQ is nearly parallel to $u = 0$, the 'distance' between P and Q can be only slightly greater than the difference between their v- (or w-) coordinates.

4. Convergence of the process. In view of the foregoing remarks, the sequence of points defined by (2) will converge to the desired solution if and only if

$$\lim_{n\to\infty} \| R_n \| = \lim_{n\to\infty} \| S_n \| = \lim_{n\to\infty} \| T_n \| = 0. \tag{18}$$

We shall prove the following result.

Theorem. Under the assumptions stated previously, the sequence defined by (2) will converge to the solution at O if the initial points R_1, S_1, T_1, satisfy the conditions

$$\max (\| R_1 \|, \| S_1 \|, \| T_1 \|) < \min \left(\frac{1}{20A}, \tfrac{3}{4}K\right), \tag{19}$$

where A and K are the constants appearing in (12), and

$$\max [r_u(S_1, T_1), r_v(R_1, T_1), r_w(R_1, S_1)] < \frac{1}{10}. \tag{20}$$

Geometrically, the conditions (19) and (20) mean that the points R_1, S_1, T_1 lie within a specified neighborhood of O and that the sides of the triangle $R_1S_1T_1$ are approximately parallel to the lines $u = 0$, $v = 0$, $w = 0$. As will become evident, other combinations of constants could be used in (19) and (20); the important point is that *some* sufficient set of bounds can be given.

Writing $M_n = \max (\| R_n \|, \| S_n \|, \| T_n \|)$ for $n = 1, 2, \cdots$, we see that to establish the theorem it is sufficient to show that the stated hypotheses imply (a) that there is a constant $\rho < 1$ such that $M_2 < \rho M_1$, and (*b*) that (20) is satisfied with R_1, S_1, T_1 replaced by R_2, S_2, T_2. For if (a) and (b) hold, the same argument can be used to show that $M_3 < \rho M_2$, $M_4 < \rho M_3$, etc., from which (18) follows immediately.

As one would expect, the proof requires repeated use of the inequalities (12), (16), and (17). For (12) to be applicable, the points involved must have norms $<K$, as is true of the initial points R_1, S_1, T_1 by hypothesis. For (16) and (17) to be applicable, directional conditions analogous to (20) must be satisfied. Defining

$$\begin{aligned} M_{11} &= \max (\| R_1 \|, \| S_1' \|, \| T_1' \|), \\ M_{12} &= \max (\| R_1' \|, \| S_1' \|, \| T_2 \|), \end{aligned} \tag{21}$$

we see that these requirements may be met by establishing the inequalities

$$M_{11} < K, \qquad r_u(S_1', T_1') < \frac{1}{10}, \qquad M_{12} < K, \qquad r_v(R_1'T_2) < \frac{1}{10}. \tag{22}$$

An outline of the proof will now be presented, followed by a more detailed discussion of some specific points.

Making use of the properties of the initial points, we begin by deriving the results

$$\max [| u(S_1') |, | u(T_1') |] \le \| R_1 \| \frac{21AM_1}{10 - 11AM_1} \le 0.111 \| R_1 \|, \tag{23}$$

and

$$\max [| w(S_1') |, | v(T_1') |] \le \| R_1 \| \left[\frac{11}{10} + \frac{21AM_1}{10(10 - 11AM_1)}\right] \le 1.11 \| R_1 \|, \tag{24}$$

from which follow, in view of (14), the relations

$$M_{11} \leq 1.22 \, || \, R_1 \, || \leq 1.22 M_1 < .92K, \qquad AM_{11} < .0611. \tag{25}$$

The next step is to establish the inequality

$$r_u(S_1'T_1') \leq \frac{11AM_1}{9 - 11AM_1} < .0651. \tag{26}$$

The results (25) and (26) permit the relations (12), (16), and (17) to be applied to the triangle $R_1S_1'T_1'$, in line with the reasoning outlined earlier.

In an analogous manner, one obtains the further inequalities

$$M_{12} \leq 1.251 \, || \, R_1 \, || < .94K, \qquad AM_{12} < .0626, \tag{27}$$

and

$$r_v(R_1'T_2) < \frac{11AM_{11}}{9 - 11AM_{11}} < .0807, \tag{28}$$

thus showing that the triangle $R_1'S_1'T_2'$ also has the desired properties.

Since T_2 would coincide with the solution point if the functions f and g were linear, one would expect $|| \, T_2 \, ||$ to be sharply bounded. In fact, one can show that

$$\begin{aligned} | \, u(T_2) \, | &\leq | \, u(T_1') \, | + \frac{r_u(S_1'T_1') \, || \, T_1' \, || \, [1 + A \, || \, T_1' \, ||]}{1 - AM_{11}[1 + r_u(S_1'T_1')]} \\ &\leq .111 \, || \, R_1 \, || + || \, T_1' \, || \, \frac{.0651(1 + AM_{11})}{1 - 1.0651AM_{11}} \leq .201 \, || \, R_1 \, || \end{aligned} \tag{29}$$

and

$$| \, v(T_2) \, | \leq || \, S_1' \, || \, \frac{21AM_{11}}{10 - 11AM_{11}} \leq .168 \, || \, R_1 \, ||, \tag{30}$$

so that by (14) we have

$$|| \, T_2 \, || \leq .369 \, || \, R_1 \, ||. \tag{31}$$

Identical arguments, applied to the triangle $R_2S_2T_2$, yield the inequalities

$$|| \, R_2 \, || \leq .254 \, || \, R_1 \, ||, \qquad || \, S_2 \, || \leq .282 \, || \, R_1 \, ||, \tag{32}$$

which together with (31) imply the relation

$$M_2 \leq .369 M_1 \, . \tag{33}$$

Thus assertion (a), stated at the outset, is verified.

By a further repetition of previous arguments, one obtains

$$r_w(R_2S_2) \leq \frac{11AM_{12}}{9 - 11AM_{12}} < .0829. \tag{34}$$

Clearly (26), (28), and (34) constitute a verification of assertion (b). Thus the proof is completed.

To a large extent, the arguments used in establishing successive steps of the preceding outline are variations on a single theme. Accordingly, only one or two steps will be considered in detail.

To begin with the proof of (23), we have in view of (2), (3), and (6)

$$u(S_1') = \frac{u(R_1)f(S_1) - u(S_1)f(R_1)}{f(S_1) - f(R_1)} = \frac{u(R_1)F(S_1) - u(S_1)F(R_1)}{u(S_1) - u(R_1) + F(S_1) - F(R_1)}$$

$$= \frac{u(R_1)\dfrac{F(S_1) - F(R_1)}{u(S_1) - u(R_1)} - F(R_1)}{1 + \dfrac{F(S_1) - F(R_1)}{u(S_1) - u(R_1)}}.$$

By (12), (17), (19), (20), and (21) we obtain

$$\left|\frac{F(S_1) - F(R_1)}{u(S_1) - u(R_1)}\right| = \frac{F(S_1) - F(R_1)\,|}{||\, S_1 - R_1\,||}\cdot\frac{||\, S_1 - R_1\,||}{|\,u(S_1) - u(R_1)\,|}$$

$$\le A[\max(||\,R_1\,||, ||\,S_1\,||)][1 + r_w(R_1S_1)] \qquad (35)$$

$$\le \frac{11}{10}AM_1\,.$$

Combining these results with (11), (13), and (19) gives us

$$|\,u(S_1')\,| \le \frac{|\,u(R_1)\,|\,\dfrac{11}{10}AM_1 + A\,||\,R_1\,||^2}{1 - \dfrac{11}{10}AM_1} \le ||\,R_1\,||\,\frac{21AM_1}{10 - 11AM_1} \le 0.111\,||\,R_1\,||,$$

and repeating the argument with S_1' replaced by T_1' yields (23).

The proof of (24) starts with the decomposition

$$w(S_1') = w(R_1) + \frac{\dfrac{w(R_1) - w(S_1)}{u(S_1) - u(R_1)}[u(R_1) + F(R_1)]}{1 + \dfrac{F(S_1) - F(R_1)}{u(R_1) - u(S_1)}}, \qquad (36)$$

and proceeds in much the same way. The proof of (26) likewise involves similar arguments, starting with the identity

$$\frac{u(S_1') - u(T_1')}{v(S_1') - v(T_1')} = \frac{\begin{vmatrix} u(R_1) - u(T_1) & f(R_1) - f(T_1) \\ u(S_1) - u(T_1) & f(S_1') - f(T_1) \end{vmatrix}}{\begin{vmatrix} v(R_1) - v(T_1) & f(R_1) - f(T_1) \\ v(S_1) - v(T_1) & f(S_1) - f(T_1) \end{vmatrix}}. \qquad (37)$$

The proofs of (27) and (28) require the use of (25) and (26), but involve no new ideas. The proof of (29) begins with a step analogous to (36), and uses the bound on $|\,u(T_1')\,|$ given by (23). The remaining steps fall into the same pattern.

Corollary. Under the hypotheses of the preceding theorem, convergence is of the second order at least.

This result is easily established. Referring to (23) and replacing AM_1 by its bound 0.05 in the denominator, we obtain

$$|\,u(S_1')\,| \le 2.22AM_1^2\,;$$

likewise,

$$|\,u(T_1')\,| \le 2.22AM_1^2\,.$$

Similarly, from (26)

$$r_u(S_1'T_1') < 1.30AM_1 .$$

Substituting these results in (29) and taking account of (25) and (26) yields

$$| u(T_2) | \leq 2.22AM_1^2 + 1.22M_1 \cdot \frac{1.30AM_1 \,(1.0611)}{1 - 1.0651\,(.0611)}$$

$$= 4.02AM_1^2 .$$

Since from (30)

$$| v(T_2) | \leq 2.17AM_1^2 ,$$

we get

$$|| T_2 || \leq 6.19AM_1^2 .$$

Similarly,

$$|| R_2 || \leq 3.61AM_1^2 , \qquad || S_2 || \leq 4.48AM_1^2 .$$

Thus $M_2 \leq 6.19AM_1^2$, which establishes the corollary.

Actually the convergence appears to be slightly faster than second-order, since $w(R_2)$ and $w(S_2)$ are of the order of M_1^3 .

5. Numerical example. To illustrate the application of the foregoing ideas, we may consider the set of equations

$$f(x, y) \equiv x^2 - 4y = 0, \qquad g(x, y) \equiv y^2 - 2x + 4y = 0, \tag{38}$$

to which we may add

$$h(x, y) \equiv -f(x, y) - g(x, y) \equiv -x^2 - y^2 + 2x = 0. \tag{39}$$

One of the common solutions of these equations is (0, 0). Taking for starting points $R_1(0, 1)$, $S_1(1, -2)$, $T_1(-1, -1)$, and applying (2) yields the sequence of points listed in the table and shown on Figs. 1, 2, and 3. The accelerating character of the convergence is evident, as is the tendency of the points to form similar triangles.

One aspect of the computation deserves further discussion. The point R_2 , which is conceptually an approximation to the point of intersection of the line $R_1'T_2$ and the curve $h = 0$, happens to fall very close to the curve $f = 0$. In consequence the points R_2 , S_2' , T_2' are much closer to each other than to the solution point, and a considerable extrapolation occurs in determining the points R_2' and T_3 . While no difficulty arises in this example, it is apparent that if R_2 had actually fallen on $f = 0$, the points S_2' and T_2' would have coincided with R_2 , and R_2' and T_3 would have been undetermined; in a practical computation, where round-off errors must be allowed for, something less that exact coincidence could cause serious trouble. A number of expedients may be imagined for meeting this problem, and are now under consideration.

From (38) and (39) we find readily

$$u = -4y, \qquad v = -2x + 4y, \qquad w = 2x, \tag{40}$$

and

$$F = x^2 = \tfrac{1}{4}w^2, \qquad G = y^2 = \frac{1}{16}u^2, \qquad H = -x^2 - y^2 = -\tfrac{1}{4}w^2 - \frac{1}{16}u^2.$$

TABLE

Example of Iterative Solution

	x	y	f	g	h
T_1	−1.0	−1.0	5.0		
S_1	1.0	−2.0	9.0		
R_1	0.0	1.0	−4.0	5.0	
T_1'	−.444	.111		1.344	
S_1'	.308	.077		−.302	.515
R_1'	.290	.130			.479
T_2	.170	.083	−.303		.304
S_2	−.0290	.0917	−.366		
R_2	−.0389	.0012	−.00329	.0826	
T_2'	−.0412	.00030		.0836	
S_2'	−.0390	.00038		.0795	−.0795
R_2'	−.0413	−.0206			−.0847
T_3	.00388	.00194	−.00774		.00774
S_3	$.0_4760$	.00180	−.00720		
R_3	$.0_4978$	$.0_4497$	$-.0_3199$		
T_3'	$-.0_5190$	$-.0_797$			
S_3'	$.0_4984$	$-.0_723$			

Since for any values of w_1 and w_2 we have

$$| \tfrac{1}{4}w_1^2 - \tfrac{1}{4}w_2^2 | = \tfrac{1}{4} | w_1 - w_2 | | w_1 + w_2 | \leq \tfrac{1}{2} | w_1 - w_2 | \max (| w_1 |, | w_2 |),$$

we conclude, referring to (10) and (11), that for any two points P, Q

$$| F(P) - F(Q) | \leq \tfrac{1}{2} || P - Q || \max (|| P ||, || Q ||).$$

Similarly, we find

$$| G(P) - G(Q) | \leq \tfrac{1}{8} || P - Q || \max (|| P ||, || Q ||),$$

and

$$| H(P) - H(Q) | \leq \tfrac{5}{8} || P - Q || \max (|| P ||, || Q ||).$$

Comparing these statements with (12), we see that the requirement (19) becomes

$$\max (|| R_1 ||, || S_1 ||, || T_1 ||) < \frac{2}{25},$$

or, if x, y are the coordinates of any one of the three points R_1, S_1, T_1,

$$\max (| 2x |, | 4y |, | 2x - 4y |) < \frac{2}{25}.$$

Substituting from (40) into (15), we can place limits on the directions of the lines R_1S_1, R_1T_1, and S_1T_1, corresponding to (20); for example, we find that the bounds on the slope of S_1T_1 are $-1/20$ and $1/21$.

It is apparent that the bounds suggested by the theorem proved earlier are much

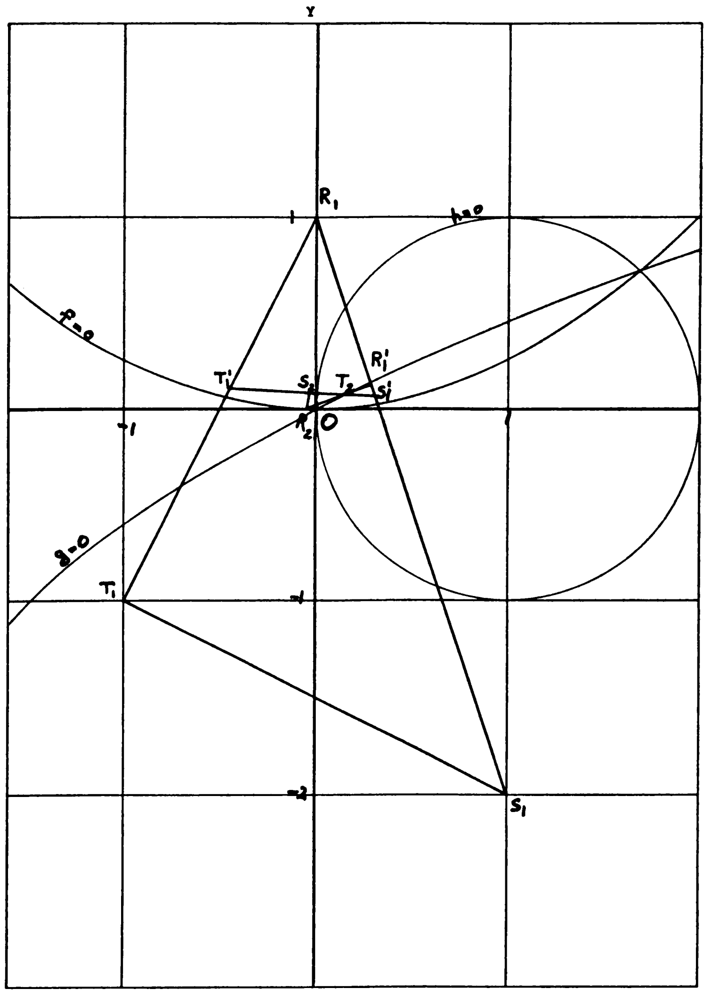
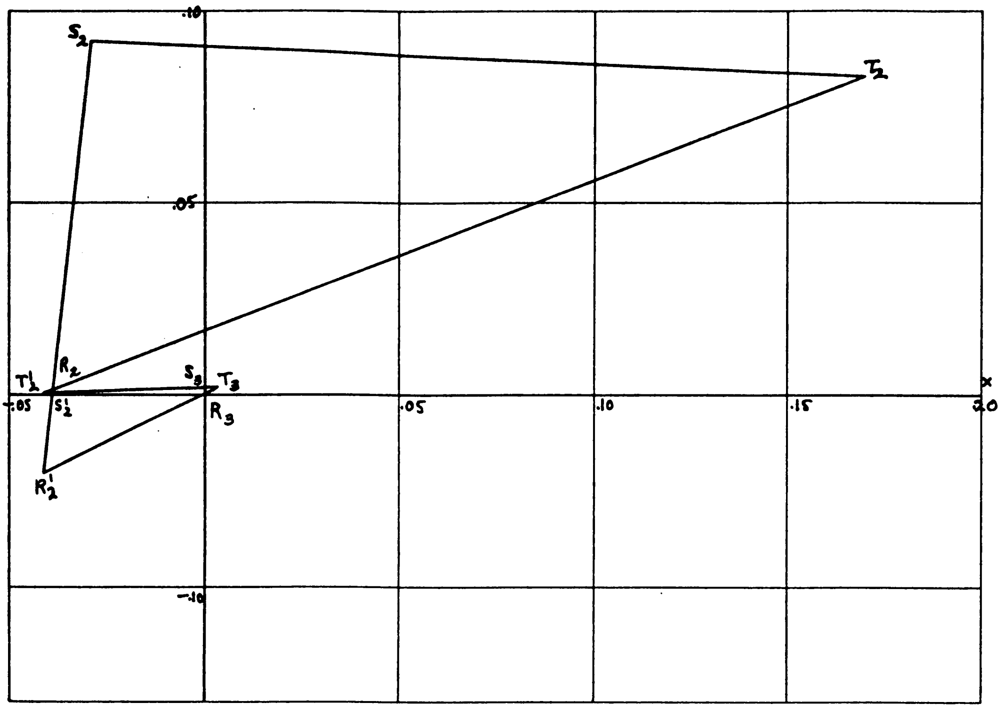


Fig. 1

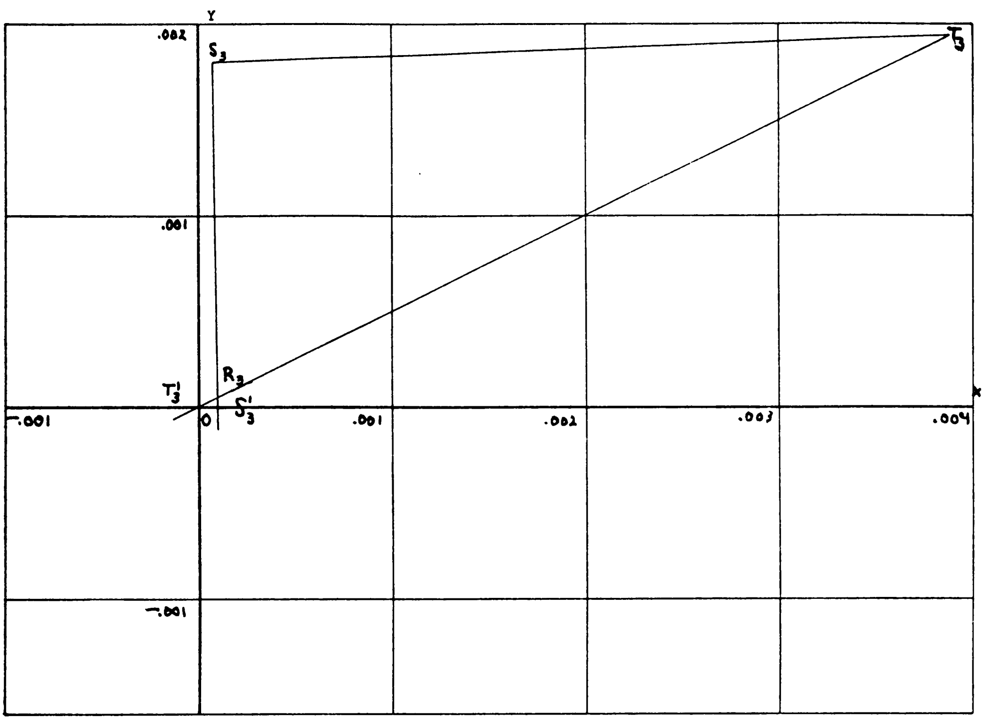
too conservative in this case. However, the result does show that the process will continue to converge once the solution has been closely approximated, and indicates the character of the sequence of iterates.

6. Concluding remarks. The foregoing example indicates that convergence may take place in practice under much less stringent conditions than were postulated in the statement of the theorem. Apart from the possibility of a more refined analysis, one must recall that in the course of the proof we necessarily assumed that each quantity appearing took on its most unfavorable value at each step, which would be highly improbable in an actual example. On the other hand, it appears that some conditions of the type imposed are necessary for convergence to the desired solution. For if the location of the initial points were uncontrolled, the iterative process might converge to another solution of the system, while their orientation must be restricted to keep the denominators in (2) away from zero.



.10
.05
-.10
-.05
.05
.10
.15
.20
x
S_2
T_2
R_2
T_2'
S_2'
S_3
T_3
R_3
R_2'

FIG. 2



Y
.002
.001
-.001
-.001
0
.001
.002
.003
.004
x
S_3
T_3
R_3
T_3'
S_3'

FIG. 3

The iterative process may clearly be generalized to apply to systems of n equations; presumably the same is true of the convergence properties. Moreover, we may note that the foregoing discussion is applicable to the finding of complex as well as real solutions.

As remarked in Sec. 2, the method of solution discussed in this paper is only one of a variety of two-point procedures that may be constructed. Some of these have been tested successfully on examples, but no sufficient conditions for convergence have been established.

References

1. K. H. Bachmann, *Der Konvergenzgrad bei iterativer Lösung von Gleichungen durch inverse Interpolation*, Z. Angew. Math. Mech. **34**, 282-3 (1956)
2. D. D. Wall, *The order of an iteration formula*, M. T. A. C., **10**, 167-8 (1956)
3. E. W. Purcell, *The vector method of solving simultaneous linear equations*, J. Math. Phys. **32**, 180-3 (1956)