# ON THE AUTOMATIC CALCULATION OF SOLUTIONS - DIFFERENT FROM THOSE PREVIOUSLY OBTAINED - OF NONLINEAR SYSTEMS OF EQUATIONS

K.M. Brown, W.B. Gearhart, and
H.A. Hall

TECHNICAL REPORT

No. 69-42

August 1969

Department of Computer Science
Cornell University
Ithaca, New York 14850

# ON THE AUTOMATIC CALCULATION OF SOLUTIONS -DIFFERENT FROM THOSE PREVIOUSLY OBTAINED- OF NONLINEAR SYSTEMS OF EQUATIONS


## ABSTRACT

Given a system of N nonlinear (algebraic or transcendental) real equations in N real unknowns, there exist a variety of numerical methods which obtain solutions of these equations. This paper presents two methods which are used to find further simple solutions - in addition to those already known a priori or from an earlier calculation. These methods have the advantage of keeping away from solutions previously calculated, saving the computer user the wasted effort entailed in converging to already known, perhaps uninteresting solutions points. The technique can also be used in avoiding previously found extreme points in function minimization. Many problems have "magnetic zeros", zeros which are converged to almost regardless of the starting guesses used. These magnetic zeros often mask out the zeros of real interest. The methods discussed are particularly effective in avoiding convergence to such magnetic zeros. Results of computer experiments are presented.

## 1. FUNCTIONS OF ONE VARIABLE.

Let $f$ be a twice continuously differentiable real valued function of the real variable $x$. (Most of the results presented apply equally well to analytic functions of a complex variable.) Let $r$ be a simple solution of the equation $f(x) = 0$; i.e., $f(r) = 0$ but $f'(r) \neq 0$. The function $g$ defined by

$$(1) \qquad \begin{aligned} g(x) &\equiv f(x)/(x-r) \ , \ x \neq r \\ g(r) &\equiv f'(r) \end{aligned}$$

is said to be obtained from $f$ by the process of deflation and $g$ is called the deflated function, or the function obtained from $f$ by deflating out the root $x = r$ (see Theorem 2 below).

Theorem 1.

The domains of definition of the functions $f$ and $g$ are identical.

Proof.

The only questionable point is $x = r$, however,

$$\lim_{x \to r} g(x) = f'(r) \ .$$

Theorem 2.

The function $g$ has all the zeros of the function $f$ except the zero $x = r$.

Proof.

For any zero $s \neq r$ of $f$, the definition of $g$ given in (1) above implies that $g(s) = 0$. On the other hand, $g(r) = f'(r) \neq 0$, since $r$ is simple.

If $f$ is a polynomial, then $g$ can be found explicitly, e.g., if

$$f(x) = x^3 - x^2 - 2x + 2$$

and

$$r = 1$$

has already been determined, then

$$g(x) = x^2 - 2.$$

For an arbitrary nonlinear $f$, however, it is not usually possible to obtain an explicit closed form expression for $g$. Even if $f$ is a polynomial, the best strategy to follow is to obtain any value, say, $g(a)$ by dividing the value of $f(a)$ by the value $a-r$ rather than forming the function $g(x)$ explicitly and then evaluating $g(a)$. The reason for this is that due to round-off errors, the coefficients of an explicitly formed $g$ may not be exact - leading to inaccuracies when computing further solutions.

We consider a natural extension of the above remarks, namely, the determination of the $k+1$ rst zero of $f$ after $k$, zeros have already been found. All zeros are assumed to be simple. The following deflation scheme

satisfies two analogues of Theorems 1 and 2 (for the k zeros case).

$$g(x) \equiv \frac{f(x)}{\prod_{i=1}^{k}(x-r_i)} \quad , \quad x \neq r_i, \; i = 1, \ldots, k,$$

(2)

$$g(r_i) \equiv \frac{f'(r_i)}{\prod_{j \neq i}(r_i - r_j)} \qquad i = 1, \ldots, k \; .$$

A popular technique for finding the zeros of a nonlinear function is given by Newton's Method:

$$(3) \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad , \quad n = 0,1,2,\ldots \quad .$$

In order to use Newton's Method for the function g given in (2), it at first seems necessary to compute $g(x_n)$ and $g'(x_n)$; however, as is known in the literature (see, e.g., [3,p.78]),

$$(4) \quad \frac{g'(x)}{g(x)} = \frac{f'(x)}{f(x)} - \sum_{i=1}^{k} \frac{1}{(x-r_i)} \quad ,$$

and the reciprocal of ratio given in (4) is precisely the value needed when applying Newton's Method to the deflated function g . The need for explicitly computing $g(x_n)$ and $g'(x_n)$ is thus circumvented. The result given in

(4) is established by writing

$$(5)\quad f(x) = g(x) \prod_{i=1}^{k} (x - r_i) ,$$

differentiating to get

$$(6)\quad f'(x) = g'(x) \prod_{i=1}^{k} (x-r_i) + g(x) \sum_{i=1}^{k} \prod_{j\neq i} (x-r_j),$$

dividing (6) by (5) and transposing.

## 2. NONLINEAR SYSTEMS OF EQUATIONS.

Let us now consider a system of $N$ nonlinear (not necessarily algebraic) real equations in $N$ real unknowns:

$$(7)\quad \begin{aligned} &f_1(x_1, x_2, \ldots, x_N) = 0 \\ &f_2(x_1, x_2, \ldots, x_N) = 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &f_N(x_1, x_2, \ldots, x_N) = 0 \end{aligned}$$

or in vector form $\underline{f}(\underline{x}) = \underline{0}$. We assume that the components of $\underline{f}$ are twice continuously differentiable with respect to all argument positions - a conditions needed for the convergence of Newton's Method. Let $\underline{r}_i$ denote the $i$th solution vector found which satisfies (7).

We now give two classes of deflations for nonlinear systems. The first which we call norm deflation is given

How about

(1)

Define $\Pi_k$ $k \times i$ matrix whose $i^{th}$ row

is $x - r_i$

Define $g(x)$ by

$$g_i(x) = \left( \frac{f_i(x)}{x_j - r_j} \right)_{j=1,\dots,n}$$

then $g_i(r) = \nabla f_i(r)$

by

$$(8)\quad \underline{g}(\underline{x}) \equiv \frac{1}{\prod_{i=1}^{k} ||\underline{x}-\underline{r}_i||}\ A\,\underline{f}(\underline{x}),\ \underline{x}\neq\underline{r}_i,\ i=1,\ldots,k\ .$$

Here $||\cdot||$ denotes any vector norm and $A$ is any non-singular matrix on $E^N$, N-dimensional Euclidean space. In the numerical results reported in Section 3 we have used $A = I$ the identity matrix; sagacious choices of $A$ for purposes of accelerating convergence or pre-conditioning the Jacobian matrix of the system are being investigated.

The second class of methods, called <u>inner product deflation</u>, is defined as follows when deflating out one root, $\underline{r}$ :

$$(9)\quad \underline{g}(\underline{x}) \equiv M\cdot\underline{f}(\underline{x});$$

here $M$ is a diagonal matrix whose <u>j</u>th diagonal element is given by

$$m_{jj} = (\langle \underline{a}_j,\ \underline{x} - \underline{r} \rangle)^{-1}$$

where $\langle,\rangle$ denotes the Euclidean inner product and $\{\underline{a}_j\}$ is a linearly independent set of vectors in $E^N$. For $k$ roots we have

$$(10)\quad \underline{g}(\underline{x}) = M_k M_{k-1} \cdots M_1\cdot\underline{f}(\underline{x})$$

with each $M_i$ defined relative to a particular root $\underline{r}_i$ and

sequence $(\underline{a}_j^i)$ as it was defined above. If for each root $\underline{r}_i$ found, we set

$$H_i(\underline{a}_j) \equiv \{\underline{u} \in E^N : \langle \underline{a}_j^i , \underline{u} \rangle = 0\} \text{ and}$$

$$D(\underline{r}_i) \equiv E^N - \bigcup_{j=1}^{N} [H_i(\underline{a}_j) + \{\underline{r}_i\}] ,$$

then the domain of definition of (10) is given by

$$D \equiv \bigcap_{i=1}^{k} D(\underline{r}_i) .$$

In our numerical investigations we chose

$$(11) \quad \underline{a}_i = \text{grad } f_i\big|_{\underline{x}=\underline{r}} \equiv f_i'(\underline{r}) .$$

We shall refer to this choice as <u>gradient deflation</u>. From (10), the explicit componentwise formulation for gradient deflation is seen to be

$$g_i^1(\underline{x}) \equiv f_i(\underline{x})$$

$$g_i^{m+1}(\underline{x}) \equiv \frac{g_i^m(\underline{x})}{\langle \text{grad } g_i^m(\underline{r}_m), \underline{x} - \underline{r}_m \rangle} , \quad m=1,\ldots,k$$

$$= \frac{f_i(\underline{x})}{\prod_{m=1}^{k} \langle \text{grad } g_i^m(\underline{r}_m), \underline{x} - \underline{r}_m \rangle} .$$

One of our objectives in defining gradient deflation was to be able to use Newton's Method for the location of the further zeros of $f$ without having to evaluate the Jacobian matrix of $g$ (see (10)) explicitly; this

would be analogous to the savings effected by (4) for Newton's Method for the one-dimensional case. Now Newton's Method in $E^N$ is given as the iteration

$$(12)\quad \underline{x}_{n+1} = \underline{x}_n - [J_{\underline{f}}(\underline{x}_n)]^{-1} \cdot \underline{f}(\underline{x}_n),\ n=0,1,\ldots,$$

where $J_{\underline{f}}(\underline{x}_n)$ denotes the Jacobian matrix of the function $\underline{f}$, defined by (7), evaluated at $\underline{x}_n$. Let $\underline{r}$ denote a zero of $\underline{f}$ and consider the function $\underline{g}$ determined by gradient deflation relative to one root and given componentwise from (9) and (11) by

$$g_i(\underline{x}) = f_i(\underline{x})/f_i'(\underline{r})\cdot(\underline{x}-\underline{r}) .$$

We seek an expression for $[J_{\underline{g}}(\underline{x})]^{-1}\,\underline{g}(\underline{x})$ which does not involve the Jacobian matrix of $\underline{g}$. Now

$$g_i'(\underline{x}) = \frac{1}{f_i'(\underline{r})(\underline{x}-\underline{r})}[f_i'(\underline{x}) - g_i(\underline{x})f_i'(\underline{r})] .$$

Let us set

$$D(\underline{r},\underline{x}) = \operatorname{diag}\{f_1'(\underline{r})(\underline{x}-\underline{r}),\ \ldots,\ f_N'(\underline{r})(\underline{x}-\underline{r})\}$$

and

$$M(\underline{r},\underline{x}) = \operatorname{diag}\{g_1(\underline{x}),\ \ldots,\ g_N(\underline{x})\} .$$

Then

$$J_{\underline{g}}(\underline{x}) = D^{-1}(\underline{r},\underline{x})[J_{\underline{f}}(\underline{x}) - M(\underline{r},\underline{x})\cdot J_{\underline{f}}^{-1}(\underline{r})\ ]$$

and therefore

$$\begin{aligned} J_{\underline{g}}^{-1}(\underline{x})\cdot\underline{g}(\underline{x}) &= [J_{\underline{f}}(\underline{x}) - M(\underline{r},\underline{x})J_{\underline{f}}(\underline{r})]^{-1}D(\underline{r},\underline{x})\cdot\underline{g}(\underline{x}) \\ &= [I - J_{\underline{f}}^{-1}(\underline{x})M(\underline{r},\underline{x})J_{\underline{f}}(\underline{r})]^{-1}J_{\underline{f}}(\underline{x})\cdot\underline{f}(\underline{x}) \end{aligned}$$

as required. This last expression is much too unwieldy for practical computations however - particularly when we observe that only one root has been deflated out. In computing the gradient vectors required, it is much more convenient to use a finite difference approximation as follows:

$$\frac{\partial g_i}{\partial x_j} \approx \frac{g_i(\underline{x} + h\underline{e}_j) - g_i(\underline{x})}{h},$$

where $\underline{e}_j$ is the $j$th unit vector and $h$ is a small constant.

REMARK.

The deflation methods defined in this section for $E^N$ have the advantage of keeping the iterates away from previously determined roots; however, unlike the one dimensional case, there is no "nice" way to define $\underline{q}$ on the sets of points excluded from the domains of definition in (8) and (9). The basic difficulty is that

$$\lim_{\underline{x} \to \underline{r}} \underline{q}(\underline{x})$$

can assume many different values depending upon the direction in which $\underline{x} \to \underline{r}$. Research is being conducted on what are good choices for $\underline{q}(\underline{r})$ relative to the iterative method being used; e.g., when using Newton's Method a good choice numerically for $\underline{q}(\underline{r})$ seems to be the value given by $J_{\underline{f}}^{-1}(\underline{r}) \cdot \underline{u}$, where $\underline{u}$ is chosen to get a unit

vector in the direction of $(\underline{x}_n - \underline{r})$ and where $\underline{x}_n$ denotes the most recent iterate found.

## 3. NUMERICAL RESULTS.

Two basic iterative methods of solution in conjunction with three types of deflation were tested on various nonlinear systems. The iterative methods used were

1) the discretized form of Newton's Method and
2) a discretized form [1] of a quadratically convergent method due to Brown [2].

The deflation techniques were

1) norm deflation with the uniform norm,
2) norm deflation with the Euclidean norm - see (8), and
3) gradient deflation;

deflations 1) and 2) will hereafter be referred to as $\ell_\infty$ deflation and $\ell_2$ deflation, respectively. The strategy of all computations was as follows: given a nonlinear system and a starting guess $\underline{x}_0$ one of the solution techniques (discretized Newton's or Brown's) was used to locate a first zero. Once accomplished one of the deflation techniques was then used with the same method beginning with the same starting guess $\underline{x}_0$ to find a second zero. If a second zero was found, deflation

was performed again, and an attempt was made to find a third zero. This process was continued, always starting with the same $\underline{x}_0$, until:

(a) all zeros of the system (or the maximum number requested by the programmer) had been found;

(b) the process diverged to "infinity";

(c) the maximum number of iterations allowed was exceeded; or

(d) the iterates reached a point at which the Jacobian matrix became singular.

All of the test cases were run in FORTRAN IV using double precision arithmetic on the IBM 360/65 at the Office of Computer Services of Cornell University. The convergence criterion was usually set to obtain twelve significant digits of accuracy.

Example 1. The Cubic-Parabola (see figure 1).

Consider the nonlinear system given by

$$f(x,y) = 4x^3 - 3x - y;\ g(x,y) = x^2 - y\ .$$

This system has the three zeros $\underline{r}_1 = (1,1)$, $\underline{r}_2 = (0,0)$ and $\underline{r}_3 = (-.75,.5625)$. An interesting feature of this system is the fact that relative to the discretized Newton's Method, $\underline{r}_1$ seems to be a magnetic zero which we define as a zero to which a particular method seems to converge no matter what the starting guess happens to be.

We gave the discretized Newton's Method 21 starting guesses in the square bounded by $x = \pm 1$, $y = \pm 1$. None of the starting guesses was closer to the first zero than to any of the other zeros, yet 20 out of the 21 starting guesses converged to $\underline{r}_1$! (One cannot help but feel sorry for the harried scientist whose physical intuition tells him that there must be another solution, yet whose computer printout always meanders back to $\underline{r}_1$ for a seemingly all encompassing set of starting guesses.) The presence of such a magnetic zero can often mask out the zeros of real interest. After removing this zero by deflation using Newton's Method, $\underline{r}_2$ was found in about half the cases tested, the other half yielding divergence. (The results for the three types of deflation were qualitatively the same.) In no case was $\underline{r}_3$ found with Newton's Method even though guesses correct to one significant digit were used. On the other hand using Brown's method $\underline{r}_3$ was found for some of the 21 starting guesses; this indicates a difference in the regions of convergence of the two methods and shows that the "magnetism" of a particular zero is a function of the method used as well as the system. More significant was the fact that from $\underline{x}_0 = (.8,.55)$, Brown's method with $\ell_\infty$ deflation found all three zeros of the system (in the order $\underline{r}_3$, $\underline{r}_2$, $\underline{r}_1$).

Beginning with the same starting guess, the zeros were found in the order $\underline{r}_3$, $\underline{r}_2$, D using $\ell_2$ deflation, and in the order $\underline{r}_3$, $\underline{r}_1$, D using gradient deflation, where D stands for divergence. Newton's Method with this $\underline{x}_0$ produced the sequence $\underline{r}_1$, $\underline{r}_2$, D for all three deflation techniques. The number of iterations used per zero by each method was comparable (12 or fewer).

Example 2. The Four-Cluster (see figure 2).

The system

$f(x,y) = (x-y^2)(x-\sin y)$

$g(x,y) = (\cos y - x)(y - \cos x)$

was designed to study the behavior of deflation when applied to a system which has several nearly equal, but distinct, zeros. This system has four zeros very close together in the first quadrant, and an infinite number of zeros spread elsewhere. The zeros of the four-cluster are $\underline{r}_1 \approx (.68,.82)$, $\underline{r}_2 \approx (.64,.80)$, $\underline{r}_3 \approx (.71,.79)$ and $\underline{r}_4 \approx (.69,.77)$. Experiments were run with eleven distinct starting guesses, the farthest away from any of the zeros being $(0,0)$. We noted the following results:

1. In no case were more than two members of the four-cluster found from the same starting guess, although other zeros not in the four-cluster were often found as well. Thus, although the deflation of one zero from the

four-cluster was able to keep iterates away from that zero but not away from its very close neighbors, the weight of two deflated zeros from the four-cluster was sufficient to direct all iterates away from the general area of the four-cluster.

2. Different deflating techniques again produced convergence to different zeros of the four-cluster from the same starting guess. Thus with $\underline{x}_0 = (.9,1)$, for example, Newton's Method converged to $\underline{r}_1$ undeflated, but converged to $\underline{r}_2$ using $\ell_2$ deflation, $\underline{r}_3$ using $\ell_\infty$ deflation and $\underline{r}_4$ using gradient deflation.

3. Divergence to infinity occurred rarely, while divergence, in the form of nonconvergence in a prescribed number of iterations (50) occurred in about a quarter of the cases. In most cases, however, the algorithms continued to find zeros using the various deflation techniques until the maximum number of zeros specified by the programmer (4) were found.

Example 3. The Hyperbola-Circle (see figure 3).

The system

$$f(x,y) = xy - 1$$
$$g(x,y) = x^2 + y^2 - 4 ,$$

has the four zeros $\underline{r}_1 = (.517,1.93)$, $\underline{r}_2 = (1.93,.517)$,

$\underline{r}_3$ = (-1.93,-.517) and $\underline{r}_4$ = (-.517,-1.93). This system appeared at first to be unstable since, for most starting guesses, Newton's Method diverged to infinity on the first attempt to find a zero. As it turned out, however, this system provided a major success for norm deflation (both types) in conjunction with Brown's method; for this combination found, from $\underline{x}_0$ = (0,1), all four zeros of the system in quick succession. (Each of the two deflators took about 35 iterations to find all four zeros.) The $\ell_2$ deflator found the zeros in the order $\underline{r}_1$, $\underline{r}_3$, $\underline{r}_2$, and $\underline{r}_4$, while $\ell_\infty$ deflation produced these roots in the order $\underline{r}_1$, $\underline{r}_4$, $\underline{r}_2$, and $\underline{r}_3$. The gradient deflator with Brown's method found only two zeros ($\underline{r}_1$ and $\underline{r}_3$) from this $\underline{x}_0$ as did Newton's Method ($\underline{r}_1$ and $\underline{r}_2$ with each deflator).

Example 4. The 3 × 3 System.

The zero-surface of

$$f(x,y,z) = x^2 + 2y^2 - 4$$

is an elliptic cylinder with elements parallel to the z-axis in 3-space; the zero surface of

$$g(x,y,z) = x^2 + y^2 + z - 8$$

is a circular paraboloid opening downward on the z-axis with vertex at (0,0,8); the zero surface of

$$h(x,y,z) = (x-1)^2 + (2y-\sqrt{2})^2 + (z-5)^2 - 4$$

is an ellipsoid having its center at the midpoint between the two zeros of the system which are given by $\underline{r}_1 = (0, \sqrt{2}, 6)$ and $\underline{r}_2 = (2,0,4)$. This problem was run using seven values of $\underline{x}_0$ for Newton's Method and using two of the seven for Brown's method. The farthest starting guess from a solution for both methods was (1,1,1), which is about a distance of 5.1 (in Euclidean norm) from $\underline{r}_1$. The other principle starting guess used was (1,.7,5), which is approximately midway between the two zeros (of Euclidean distance = 1.58 from each). In all cases which were run, both zeros were found in at most 19 iterations per zero, except for Newton's Method starting at (1,1,1) with $\ell_2$ deflation which failed to find the second zero in 200 iterations.

To summarize (a large number of numerical experiments), $\ell_\infty$ norm deflation, besides being the easiest to compute, was the most stable of the three types of deflation studied.

## 4. APPLICATIONS TO FUNCTION MINIMIZATION.

Given a real function of N real variables,

$$F(x_1, x_2, \ldots, x_N),$$

a technique often used to locate minima of F is to determine the zeros of the syetem of partial derivatives

obtained by differentiating F with respect to each of its argument positions. This yields a nonlinear system of equations to be solved (see (7)). The solutions of this system furnish candidates for the desired minimum points. The deflation techniques developed above may be used in avoiding such previously found minima.

## 5. CONCLUSIONS.

Deflation seems to be a useful tool to incorporate into programs which are written to solve nonlinear systems of equations, particularly when such systems possess magnetic zeros. Just as deflation seems to work better with Muller's method as opposed to Newton's method in the one dimensional case, our investigations indicate that deflation should perhaps be tied to methods other than Newton's in the higher dimensional cases as well.

## REFERENCES

[1] K.M. Brown, "Solution of Simultaneous Non-linear Equations", Comm. Assoc. Comput. Mach., Vol. 10 (1967), pp. 728-729.

[2] K.M. Brown, "A Quadratically Convergent Newton-like Method Based upon Gaussian Elimination", accepted for publication in SIAM J. Numer. Anal. (Preprint available as Department of Computer Science Technical Report No. 68-23, Cornell University, Ithaca, New York 14850).

[3] J.H. Wilkinson, Rounding Errors in Algebraic Processes, Prentice-Hall, Inc., Englewood Cliffs, 1963.

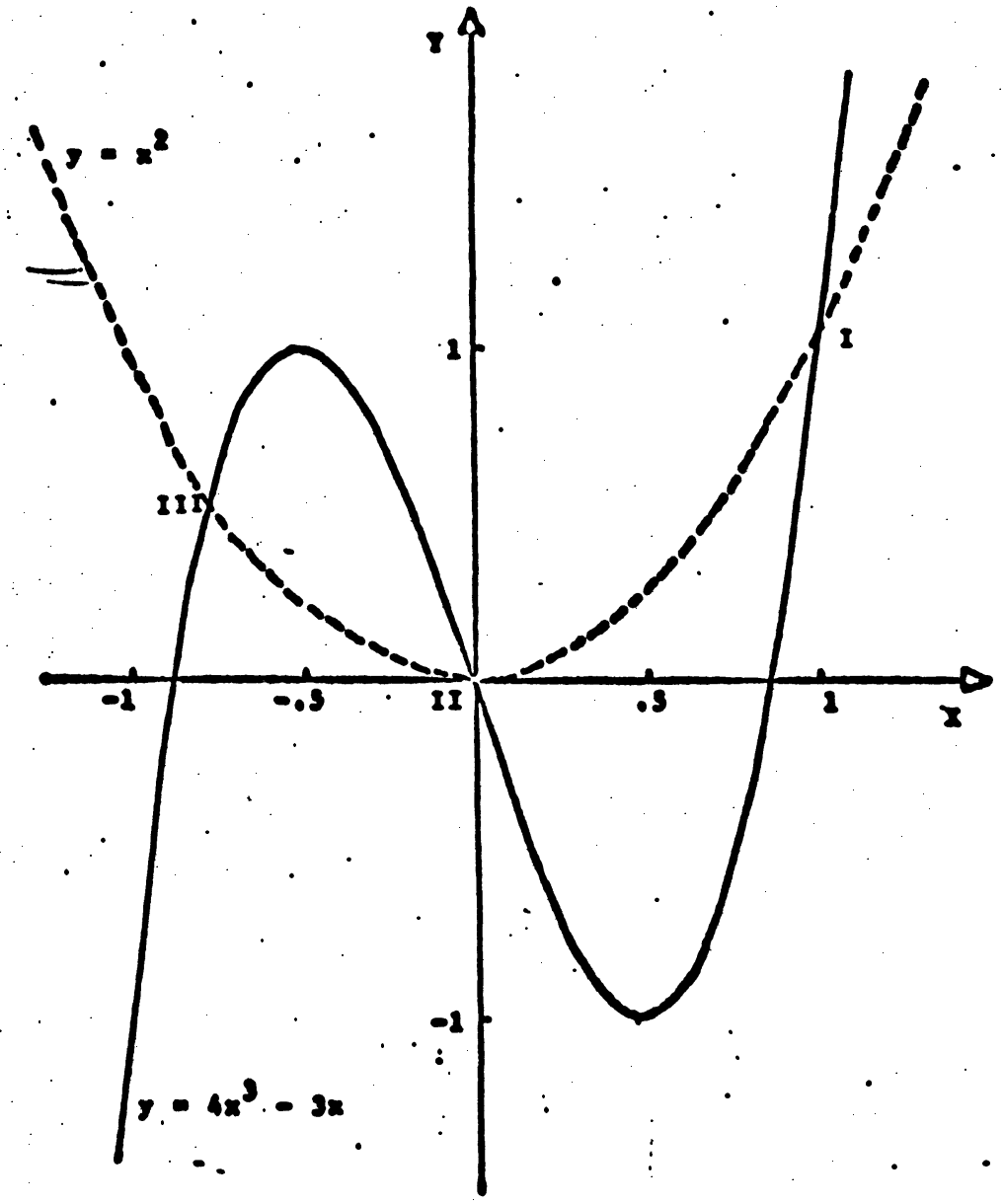


Fig. 1. Zero Curves of the Cubic-Parabola System

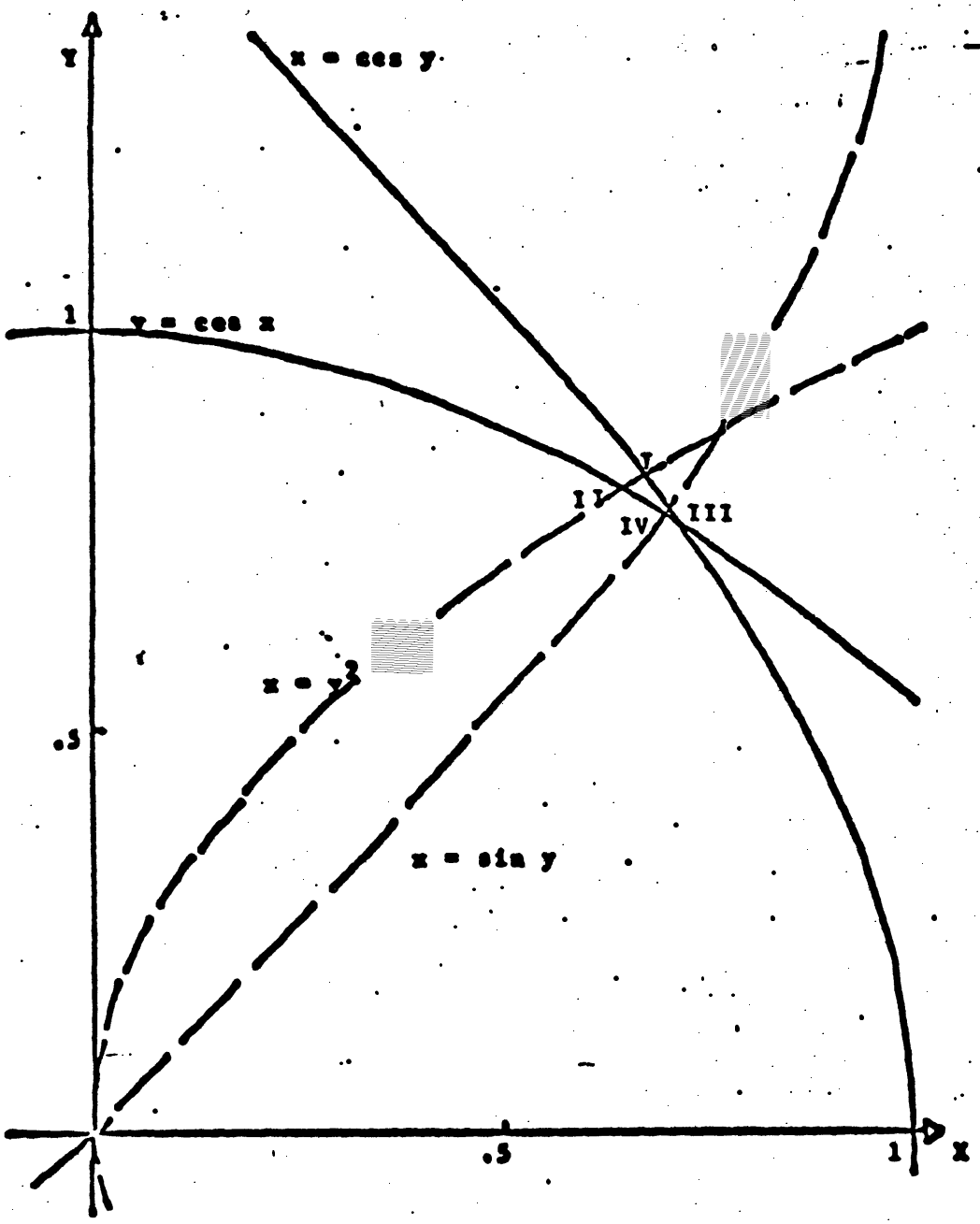


Fig. 2. Zero Curves of the Four-Cluster System

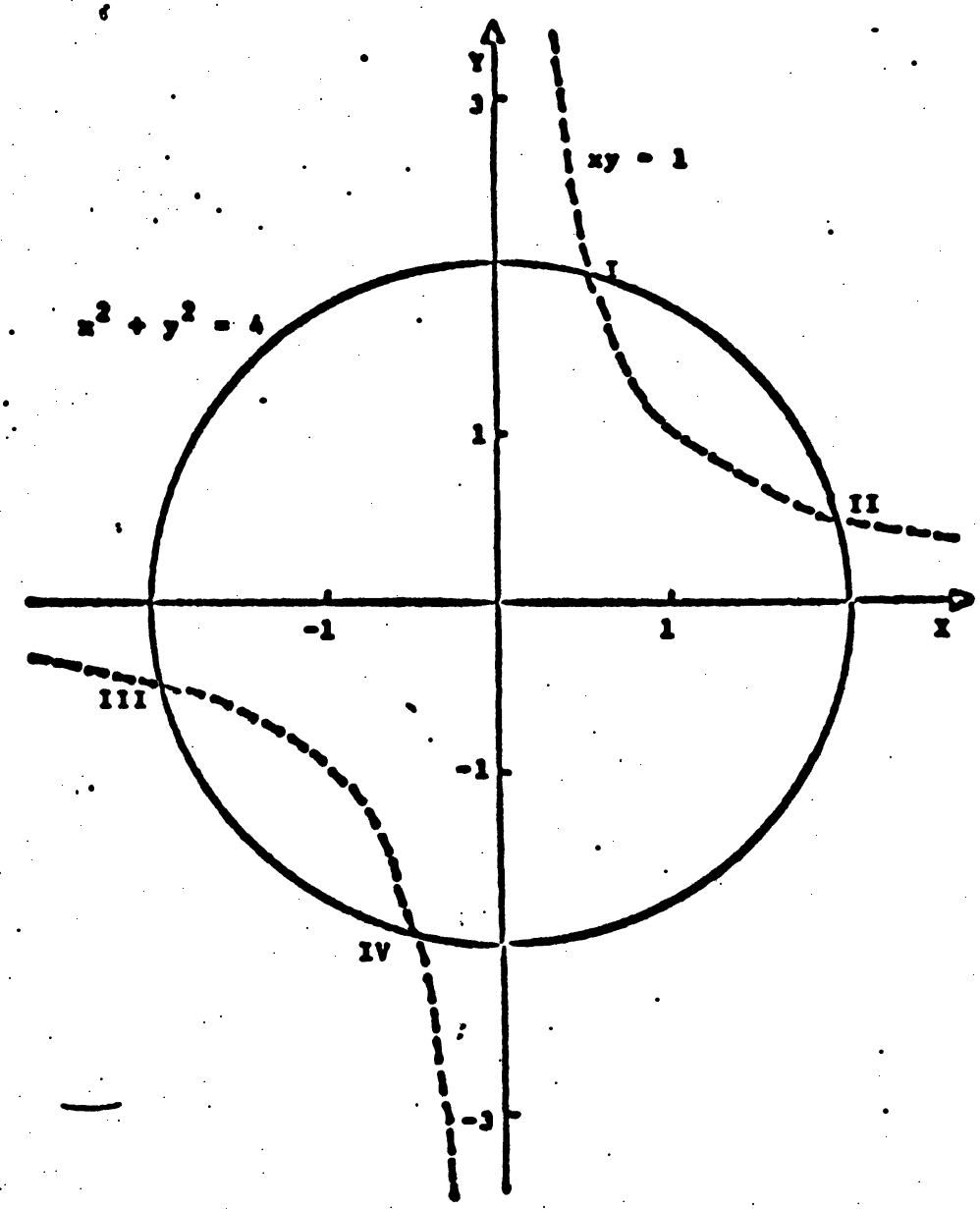


Fig. 3. Zero Curves of the Hyperbola-Circle System