General Disclaimer

One or more of the Following Statements may affect this Document

- This document has been reproduced from the best copy furnished by the organizational source. It is being released in the interest of making available as much information as possible.

- This document may contain data, which exceeds the sheet parameters. It was furnished in this condition by the organizational source and is the best copy available.

- This document may contain tone-on-tone or color graphs, charts and/or pictures, which have been reproduced in black and white.

- This document is paginated as submitted by the original source.

- Portions of this document are not fully legible due to the historical nature of some of the material. However, it is the best reproduction available from the original submission.

SQT

TWISK 231

Geometrically derived difference formulae for the numerical integration of trajectory problems

by

R J Y McLeod
Department of Mathematical Sciences
New Mexico State University, Las Cruces,
New Mexico, U S A

and

J M Sanz-Serna
Departamento de Matemáticas, Facultad
de Ciencias, Universidad del Pais Vasco,
Lejona (Vizcaya), Spain

(both authors visiting scientists at NRIMS when this work was done)

Tegniese Verslag **Technical Report**
NASIONALE NAVORSINGSINSTITUUT VIR WISKUNDIGE WETENSKAPPE
NATIONAL RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES
WNNR **CSIR**

TWISK 231. 19 p. + 13 figs., Pretoria. September 1981

TWISK 231

Geometrically derived difference formulae for the numerical integration of trajectory problems

by

R J Y McLeod
Department of Mathematical Sciences
New Mexico State University, Las Cruces,
New Mexico, U S A

and

J M Sanz-Serna
Departamento de Matemáticas, Facultad
de Ciencias, Universidad del Pais Vasco,
Lejona (Vizcaya), Spain

(both authors visiting scientists at NRIMS when this work was done)

Tegniese Verslag **Technical Report**
NASIONALE NAVORSINGSINSTITUUT VIR WISKUNDIGE WETENSKAPPE
NATIONAL RESEARCH INSTITUTE FOR MATHEMATICAL SCIENCES
WNNR **CSIR**

TWISK 231, 19 p. + 13 figs., Pretoria, September 1981

NRIMS Technical Report TWISK 231
National Research Institute for Mathematical Sciences
CSIR
P O Box 395
PRETORIA
0001

Printed in the
Republic of South Africa
by the Graphic Arts Division of the CSIR
PRETORIA

GEOMETRICALLY DERIVED DIFFERENCE FORMULAE FOR THE NUMERICAL INTEGRATION OF TRAJECTORY PROBLEMS

by

R J Y McLeod*
Department of Mathematical Sciences
New Mexico State University, Las Cruces
New Mexico 88003, U S A

and

J M Sanz-Serna*
Departamento de Matematicas, Facultad
de Ciencias, Universidad del Pais Vasco
Lejona (Vizcaya), Spain

ABSTRACT

The term 'trajectory problem' is taken to include problems that can arise, for instance, in connection with contour plotting, or in the application of continuation methods, or during phase-plane analysis. Geometrical techniques are used to construct difference methods for these problems to produce in turn explicit and implicit circularly exact formulae. Based on these formulae, a predictor-corrector method is derived which, when compared with a closely related standard method, shows improved performance. It is found that this latter method produces spurious limit cycles, and this behaviour is partly analyzed. Finally, a simple variable-step algorithm is constructed and tested.

---

*Visiting Scientists at the National Research Institute for Mathematical Sciences, CSIR, P O Box 395, Pretoria, South Africa.

1. INTRODUCTION

We consider an initial-value problem for the autonomous system of ordinary differential equations

$$\frac{d\underset{\sim}{y}}{dt} = \underset{\sim}{f}(\underset{\sim}{y}) \ , \tag{1.1}$$

where $\underset{\sim}{y}$ is a vector in $\mathbb{R}^m$.

In a number of practical applications the interest lies in obtaining the curve traced by the solution $\underset{\sim}{y}(\cdot)$ rather than in finding the actual correspondence between values of the independent variable or parameter t and points on that curve. These applications include the computation of trajectories in mechanical problems, the plotting of the phase-plane of second-order autonomous differential equations [2], and the study of solution fields of nonlinear equations [1,5]. We shall employ the term trajectory problem to refer to these cases.

By definition a trajectory problem is not altered if the independent variable in (1.1) is replaced by a new variable $u = \varphi(t)$, where φ is differentiable and monotonic. On the other hand the performance of a numerical method when applied to (1.1) depends heavily on the particular parametrization [6]. To overcome the difficulties associated with the choice of this independent variable, the following devices come easily to mind.

(i) Use of one of the coordinates, the first say, of $\underset{\sim}{y}$ as independent variable. This procedure reduces by one the dimension of the system, but suffers from the disadvantage that the integration cannot be carried beyond a point $\underset{\sim}{y}$ for which $f_1(\underset{\sim}{y}) = 0$. It should also be noted that this procedure is not invariant with respect to rotation of the axes in the y-space.

(ii) Parametrization of the curve by its unique intrinsic parameter, i.e. its arc length s. This is equivalent to replacing (1.1) by

$$\frac{dy}{ds} = \frac{1}{\|f(y)\|} f(y) =: F(y) , \qquad (1.2)$$

since now $\|dy/ds\| = 1$. (We shall here not be concerned with singular points where $f(y) = 0$.) The use of the arc length and some of its modifications has been advocated by H B Keller [4] in the context of the solution of nonlinear equations. See also [6].

For the two-dimensional case (m=2) Lambert and McLeod [2] have introduced a successful modification of the idea in (i). They use the midpoint rule by rotating locally the axes in the y plane so as to have the tangent to the solution at the latest computed point playing the role of positive direction of the independent variable. This local rotation renders their method intrinsic in the sense that it does not depend upon the orientation of the axes in the y plane. Lambert and McLeod prove their method to be circularly exact, i.e. if the trajectory is a circle all the computed points will lie on the circle, provided that the starting points do and that no round-off error is present. Laurie [3] has extended the idea of local rotation to higher-dimensional equations.

It appears to be desirable that a method should be circularly exact, as any m-dimensional curve can be approximated to second-order terms by its local circle (see Section 3).

This paper continues the study of difference schemes specifically derived for trajectory problems.

In Section 2 we present a simple geometrical way of constructing such methods.

The local accuracy of the schemes is investigated in Section 3.

In Section 4 we define a circularly exact, fixed-step predictor-corrector algorithm that is closely related to the standard predictor-corrector method comprising the mid-point and trapezoidal rules in PECE mode.

When both algorithms are tested in a number of problems the standard method is found to produce spurious limit cycles in some cases. It is proved that for a model problem the spurious cycles are local attractors.

In the final Section we present a variable-step version of the circularly exact algorithm, a version whose step control strategy is based on a Milne device. Numerical examples are given.

2. GEOMETRICAL CONSTRUCTION

We illustrate the general idea by constructing the circularly exact method of Lambert and McLeod. This is an explicit, two-step formula which computes $\underset{\sim}{y}_{n+2}$ in terms of the back points $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$ and the back slope $\underset{\sim}{f}_{n+1} := \underset{\sim}{f}(\underset{\sim}{y}_{n+1})$. We note that the points $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$ and the vector $\underset{\sim}{f}_{n+1}$ uniquely determine a circle C_n in the m-dimensional space, the circle degenerating to a straight line if $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$, $\underset{\sim}{f}_{n+1}$ are parallel. Choice of any point $\underset{\sim}{y}_{n+2}$ on C_n makes the formula circularly exact. In particular we can define $\underset{\sim}{y}_{n+2}$ to be the point on C_n such that $\|\underset{\sim}{y}_{n+2} - \underset{\sim}{y}_{n+1}\| = \|\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n\|$ (cf. Figure 1, which depicts the two-dimensional plane spanned by the points $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$ and the vector $\underset{\sim}{f}_{n+1}$).

It is clear that with this choice

$$\underset{\sim}{y}_{n+2} = \underset{\sim}{y}_n + 2\,[\underset{\sim}{F}^{\mathsf{T}}_{n+1}(\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n)]\,\underset{\sim}{F}_{n+1}, \tag{2.1}$$

where $\underset{\sim}{F}_{n+1} = \underset{\sim}{f}_{n+1}/\|\underset{\sim}{f}_{n+1}\|$, and this is precisely the Lambert-McLeod method as written by Laurie [3].

By construction the method generates points such that $\|\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n\|$ is constant. In a 'variable-step' implementation one may wish to increase or decrease the Euclidean distance between consecutive points, and this

can be achieved by changing the choice of $\underset{\sim}{y}_{n+2}$ on C_n, as will be shown in Section 5.

Turning now to the general idea, suppose that we are given a family of curves such that an individual member of the family can be determined by M linear conditions (when m = 2 three conditions determine a circle, four a parabola, five a general conic, etc....). Then M pieces of information from the back data can be used to determine a curve of the family, and any choice of the next point on this curve will yield an explicit method which is exact whenever the trajectory belongs to the given family.

This idea can also be employed to derive implicit methods. In this case the slope at the next point appears in the formula, and only M - 1 pieces of information from the back data are required. As an illustration, let us derive a circularly exact one-step method. From Figure 2 we see that when the solution is a circle, $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$ bisects the angle between the unit vectors $\underset{\sim}{F}_n$, $\underset{\sim}{F}_{n+1}$.

Therefore

$$\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n = k\tfrac{1}{2}(\underset{\sim}{F}_n + \underset{\sim}{F}_{n+1}), \tag{2.2}$$

where k is a parameter, yields the method sought for. Of course (2.2) is nothing but the trapezoidal rule applied to (1.2) with step-size k.

3. THE TRUNCATION ERROR

In this section we attempt to define the concept 'truncation error' for methods such as (2.1). In order to motivate the definition, let us consider first the formula (2.2). When this is viewed as the usual trapezoidal rule applied to (1.2), the standard procedure is to define the truncation error at a point $y(s_o)$ of the trajectory by

$$TE = \underset{\sim}{y}(s_o + k) - \underset{\sim}{y}(s_o) - \frac{k}{2}\left[\left.\frac{d\underset{\sim}{y}}{ds}\right|_{s_o} + \left.\frac{d\underset{\sim}{y}}{ds}\right|_{s_o + k}\right]. \qquad (3.1)$$

A Taylor expansion reveals that as $k \to 0$

$$TE = -\frac{1}{12} k^3 \left.\frac{d^3\underset{\sim}{y}}{ds^3}\right|_{s_o} + O(k^4) \qquad (3.2)$$

and accordingly one says that the method is of second order. We recall that if we denote by $\underset{\sim}{t}$, $\underset{\sim}{n}$, $\underset{\sim}{b}$ the local tangent, (first) normal and second normal unit vectors, respectively, the derivatives of $\underset{\sim}{y}$ w.r.t. s can be expressed as follows:

$$\dot{\underset{\sim}{y}} = \underset{\sim}{t} \qquad (3.3)$$

$$\ddot{\underset{\sim}{y}} = \kappa \underset{\sim}{n}$$

$$\dddot{\underset{\sim}{y}} = \dot{\kappa}\underset{\sim}{n} - \kappa^2 \underset{\sim}{t} + \kappa\tau\underset{\sim}{b}.$$

Here a dot represents differentiation with respect to the arc length s and κ, τ the first and second curvatures. When the curve is three-dimensional the terms binormal and torsion are often used to refer to $\underset{\sim}{b}$ and τ respectively.

From these expressions we see that in the neighbourhood of a point any m-dimensional curve can be approximated to second-order accuracy by the circle which shares its curvature, and tangent and normal vectors. When (3.3) is taken into account (3.1), (3.2) can be written as

$$TE = \underset{\sim}{y}(s_o+k) - \underset{\sim}{y}(s_o) - \frac{k}{2}[\underset{\sim}{t}(s_o) + \underset{\sim}{t}(s_o+k)] = \qquad (3.4)$$

$$= -\frac{1}{12} k^3 [\dot{\kappa}(s_o)\,\underset{\sim}{n}(s_o) - \kappa^2(s_o)\,\underset{\sim}{t}(s_o) + \kappa(s_o)\tau(s_o)\underset{\sim}{b}(s_o)] + O(k^4).$$

When the true trajectory is a circle, $\dot{\kappa},\tau \equiv 0$ and (3.4) becomes

$$TE = \frac{1}{12} k^3 \kappa^2 \underset{\sim}{t} + O(k^4). \qquad (3.5)$$

The fact that we are dealing with a circularly exact method is not apparent from (3.4). This is due to the fact that the truncation error locally measures the distance between the computed point $\underset{\sim}{y}_{n+1}$ and the exact $\underset{\sim}{y}(s_{n+1})$ (when $\underset{\sim}{y}_n = \underset{\sim}{y}(s_n)$), whilst we are interested in the distance between $\underset{\sim}{y}_{n+1}$ and the trajectory.

As an alternative we shall define the concept of reduced truncation error (RTE) which has the following property: whenever the method is exact for a family of curves in the sense of the previous section, the RTE for a trajectory on that family vanishes identically.

For the particular case of the trapezoidal rule we proceed as follows: we denote by $h = h(k)$ the Euclidean distance between $\underset{\sim}{y}_{n+1}$ and $\underset{\sim}{y}_n$ when $\underset{\sim}{y}_n = \underset{\sim}{y}(s_o)$, and then define the RTE at $\underset{\sim}{y}(s_o)$ by

$$RTE = \underset{\sim}{y}^* - \underset{\sim}{y}(s_o) - \frac{k}{2}[\underset{\sim}{t}(s_o) + \underset{\sim}{t}^*] \tag{3.6}$$

where $\underset{\sim}{y}^*$ is the point on the trajectory such that $\|\underset{\sim}{y}^* - \underset{\sim}{y}(s_o)\| = h$ and $\underset{\sim}{t}^*$ is the unit tangent vector at $\underset{\sim}{y}^*$.

Thus whenever a step of the trapezoidal rule starting from $\underset{\sim}{y}_n = \underset{\sim}{y}(s_o)$ leads to a point $\underset{\sim}{y}_{n+1}$ which lies on the trajectory, we shall have $\underset{\sim}{y}^* = \underset{\sim}{y}_{n+1}$ and hence RTE = 0.

Let us now expand the RTE (3.6) in powers of h. In order to do so we reparametrize the trajectory in the neighbourhood of $\underset{\sim}{y}(s_o)$, taking as new parameter the Euclidean distance $h(s) = \|\underset{\sim}{y}(s) - \underset{\sim}{y}(s_o)\|$. Taylor expansion of $\underset{\sim}{y}(s) - \underset{\sim}{y}(s_o)$ and use of (3.3) reveal that

$$h = (s-s_o) - \frac{\kappa^2}{24}(s-s_o)^3 + O((s-s_o)^4). \tag{3.7}$$

Now the standard rules for the differentiation of inverse and composite functions yield the following expressions for the derivatives of y w.r.t. h:

$$\underset{\sim}{dy}/dh = \underset{\sim}{t}, \tag{3.8}$$

$$d^2\underset{\sim}{y}/dh^2 = \kappa\underset{\sim}{n},$$

$$d^3\underset{\sim}{y}/dh^3 = \dot{\kappa}\underset{\sim}{n} - 3/4\ \kappa^2\underset{\sim}{t} + \kappa\tau\ \underset{\sim}{b}.$$

Analogously, for the derivatives of $\underset{\sim}{t}$ one has

$$d\underset{\sim}{t}/dh = \kappa\underset{\sim}{n}, \tag{3.9}$$

$$d^2\underset{\sim}{t}/dh^2 = \dot{\kappa}\underset{\sim}{n} - \kappa^2\underset{\sim}{t} + \kappa\tau\underset{\sim}{b}.$$

Next, we eliminate k from (3.6), noting that

$$k = \frac{2\|\underset{\sim}{y}^* - \underset{\sim}{y}(s_o)\|}{\|\underset{\sim}{t}(s_o) + \underset{\sim}{t}^*\|} = \frac{2h}{\|\underset{\sim}{t}(s_o) + \underset{\sim}{t}^*\|}\ . \tag{3.10}$$

Substituting (3.10) into (3.6) and expressing the result in terms of the parameter h, we have

$$RTE = \underset{\sim}{y}(h) - \underset{\sim}{y}(0) - \frac{2h}{\|\underset{\sim}{t}(0) + \underset{\sim}{t}(h)\|}\,[\underset{\sim}{t}(0) + \underset{\sim}{t}(h)]. \tag{3.11}$$

We now Taylor-expand, using (3.8), (3.9) to replace the derivatives of $\underset{\sim}{y}$, $\underset{\sim}{t}$, and arrive at

$$RTE = -\frac{1}{12}\,h^3[\dot{\kappa}\underset{\sim}{n} + \kappa\tau\underset{\sim}{b}] + O(h^4). \tag{3.12}$$

We note that the curvature does not appear alone in the leading terms of the RTE, in agreement with the fact that RTE = 0 if $\dot{\kappa}$, $\tau = 0$. (In fact it can be shown that the whole Taylor series for RTE does not involve terms which contain only the curvature.)

The idea we have just illustrated in the case of the trapezoidal rule can be extended to other members of the class of methods introduced in the previous section. For instance for the method (2.1) one would define

$$RTE = \underset{\sim}{y}^{**} - \underset{\sim}{y}(s_o) + 2\,[\underset{\sim}{t}^{*T}\,(\underset{\sim}{y}^* - \underset{\sim}{y}(s_o))]\,\underset{\sim}{t}^*\ , \tag{3.13}$$

where $\underline{y}^*$, $\underline{y}^{**}$ are the points on the curve such that $\|\underline{y}(s_o)-\underline{y}^*\| = \|\underline{y}^*-\underline{y}^{**}\| = h$ with h equal to the constant distance between any two consecutive points. We now find

$$\text{RTE} = \frac{h^3}{3}[\dot{\kappa}\underline{n} + \kappa\tau\underline{b}] + O(h^4). \tag{3.14}$$

This idea of an RTE can be employed to derive estimates of the global accuracy of the methods. The details will be given elsewhere.

4. A CIRCULARLY EXACT PREDICTOR-CORRECTOR METHOD

Comparison of (3.12) with (3.14) shows that the implicit circularly exact method (2.2) has a smaller error constant than the explicit method (2.1). Therefore it is reasonable to consider the idea of combining the two methods in a predictor-corrector pair. We suggest the following formulae:

$$\underline{y}^P_{n+2} = \underline{y}_n + 2\,[\underline{F}^T_{n+1}\,(\underline{y}_{n+1} - \underline{y}_n)]\,\underline{F}_{n+1}, \tag{4.1}$$

$$\underline{y}_{n+2} = \underline{y}_{n+1} + \frac{h}{\|\underline{F}_{n+1} + \underline{F}^P_{n+2}\|}\,(\underline{F}_{n+1} + \underline{F}^P_{n+2}),$$

where $h = \|\underline{y}_0 - \underline{y}_1\|$, $\underline{F}^P_{n+2} = \underline{F}(\underline{y}^P_{n+2})$.

Note that $\|\underline{y}^P_{n+2} - \underline{y}_{n+1}\| = \|\underline{y}_{n+1} - \underline{y}_n\|$ and that the step-length k of the corrector (2.2) is changed from one step to the next in order to guarantee that $\|\underline{y}_{n+2} - \underline{y}_{n+1}\| = h$.

When the trajectory is a circle and $\underline{y}_n$, $\underline{y}_{n+1}$, lie on the trajectory, the predictor yields a point $\underline{y}^P_{n+2}$ on the circle with $\|\underline{y}_{n+1} - \underline{y}^P_{n+2}\| = h$. Therefore $\underline{y}_{n+2} = \underline{y}^P_{n+2}$ and the method is circularly exact.

Formulae (4.1) were tested in several numerical examples, and in order to establish a fair comparison the following method was used:

$$\underline{y}^{p}_{n+2} = \underline{y}_n + 2k\,\underline{F}_{n+1}, \tag{4.2}$$

$$\underline{y}_{n+2} = \underline{y}_{n+1} + (k/2)\,(\underline{F}_{n+1} + \underline{F}^{p}_{n+2}),$$

i.e. the predictor-corrector method based on the mid-point and trapezoidal rules used in PECE mode. Recall that $\underline{F} = \underline{f}/|\underline{f}|$.

It should be stressed that if correction to convergence rather than the PECE mode had been used, one would have had the circularly exact method (2.2). However (4.2) is not circularly exact, as will be clear from the following discussion.

Suppose that (4.2) is applied to the two-dimensional problem

$$f_1 = -y_2 \tag{4.3}$$

$$f_2 = \ \ y_1$$

whose trajectories are circles centered at the origin.

This problem is best analyzed by means of polar coordinates. Namely let us describe each of the vectors $\underline{y}_n$ generated by (4.2) by the radius $\rho_n = |\underline{y}_n|$ and the angle α_n formed by $\underline{y}_{n-1}, \underline{y}_n$. Then, after some manipulation, it is found that $\underline{y}_{n+2}$ is obtained from $\underline{y}_{n+1}, \underline{y}_n$ by means of the formulae

$$\rho_{n+2} = (k^2 \cos^2\beta + \rho^2_{n+1} - k\rho_{n+1} \sin 2\beta)^{\frac{1}{2}}, \tag{4.4a}$$

$$\cos \alpha_{n+2} = (\rho^2_{n+1} + \rho^2_{n+2} - k^2 \cos^2\beta)/(2\,\rho_{n+1}\,\rho_{n+2}), \tag{4.4b}$$

where β is a function of k, ρ_n, α_{n+1} given by

$$\cot 2\beta = \rho_n \cos \alpha_{n+1}/(2k - \sin \alpha_{n+1}). \tag{4.5}$$

We see from (4.4a) that in general the radius ρ does not remain constant for all iterants and therefore that the method is not circularly exact.

It is useful to take this discussion further as follows. Formulae (4.4)

describe a two-step recurrence for the computation of $(\rho_{n+2}, \alpha_{n+2})$ in terms of $(\rho_{n+1}, \alpha_{n+1})$ and ρ_n. It is possible to reformulate this recurrence as one having only one step, by increasing the dimension of the vectors involved. Namely with $r_{n+1} = \rho_n$ we arrive at the recurrence

$$\begin{aligned} r_{n+2} &= \rho_{n+1}, \\ \rho_{n+2} &= R, \\ \alpha_{n+2} &= \arccos\,(\rho^2_{n+1} + R^2 - k^2\cos^2\beta)/(2\,\rho_{n+1}\,R)], \end{aligned} \tag{4.6}$$

where β, R satisfy

$$\begin{aligned} \cot 2\beta &= r_{n+1}\cos\alpha_{n+1}/(2k - r_{n+1}\sin\alpha_{n+1}) \\ R &= (k^2\cos^2\beta + \rho^2_{n+1} - k\,\rho_{n+1}\sin 2\beta)^{\frac{1}{2}}. \end{aligned} \tag{4.7}$$

Now (4.6) describes the transformation of $(r_{n+1}, \rho_{n+1}, \alpha_{n+1})$ into $(r_{n+2},\rho_{n+2},\alpha_{n+2})$. It is easily verified that $(k/2, k/2,\pi/2)$ is a fixed point of this iteration.

We conclude that if (4.2) is applied to the model system (4.3) with $\|\underset{\sim}{y}_0\| = \|\underset{\sim}{y}_1\| = k/2$ and $\underset{\sim}{y}_0$, $\underset{\sim}{y}_1$ forming an angle of $\pi/2$, then each subsequent iterant also lies on a circle of radius $k/2$ and is $\pi/2$ radians from the previous iterant. We shall use the term 'spurious limit circle' to refer to this circle of radius $k/2$.

The Jacobian matrix of the transformation (4.6) evaluated at the fixed point is found to be

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -k/6 \\ 0 & -2/k & 0 \end{bmatrix}$$

with eigenvalues $0, \pm \sqrt{3}/3$. Since these are smaller in magnitude than unity, the fixed point is a local attractor, i.e. initial vectors $\underset{\sim}{y}_0, \underset{\sim}{y}_1$ near the spurious limit circle and forming an angle near to $\pi/2$ will pro-

duce a sequence of iterants which converges to the spurious limit circle. In fact we shall see in what follows that the iterants converge to that circle even if the initial vectors are far from it.

We are now in a position to report several numerical tests on methods (4.1), (4.2). In all the examples the 'exact' trajectory was calculated employing the usual fourth-order, fourth-stage Runge-Kutta method, which also provided the additional starting value.

When using the Runge-Kutta method, a step-size one-tenth that of the predictor-corrector algorithms was taken. In the figures the points produced by the Runge-Kutta method have been joined by a continuous curve, those produced by (4.1) being indicated by circles 'O' and those produced by the trapezoidal rule indicated by crosses 'X', and joined by a broken line for additional clarity.

As a first example we consider the model problem (4.3). The initial point was (0,1) and the step 1. ('Step' means, of course, h in formulae (4.1), k in formulae (4.2).) The results have been plotted in Fig.3. Ninety-eight points were computed for each algorithm. Those corresponding to the circularly exact method fall repeatedly on the inscribed hexagon, showing numerical stability. The points corresponding to (4.2) spiral very rapidly towards the spurious limit circle, and from the sixteenth onwards lie on that circle (within the accuracy of the plot). Fig.4 corresponds to the same problem and initial condition, but the step is now 0.37. Note that the radius of the spurious limit circle has decreased, in agreement with our earlier discussion.

The second example is the system

$$f_1 = - y_2, \qquad (4.8)$$

$$f_2 = \sin y_1,$$

which is equivalent to the well-known pendulum equation.

The initial point was (0,1) and the step 0.5. The behaviour of the methods was very similar to the one we have seen in the first example. The points produced by the circularly exact method were, within the accuracy of the figure, on the exact integral curve. The solution given by the method (4.2) spiralled in and reached a limit circle of radius 0.25, far from the true orbit. This value of the radius is precisely that of the spurious limit circle for the model problem. This is no surprise as the phase-planes of (4.3), (4.8) near the origin are very similar.

The next example is the van der Pol system

$$f_1 = y_2 - .1(y_1^3 - 3\, y_1) \qquad (4.9)$$

$$f_2 = -y_1.$$

The results illustrated by Figs. 5 and 6 both refer to a step 1.5 but the starting point was (10,10) for the former and (0,1) for the latter. We see that in both instances the circularly exact method identifies correctly the limit cycle of the system, whereas the results given by the method (4.2) suggest a 'spurious' limit cycle whose diameter is roughly half the true one. Neither method does well in the descending section of the trajectory in Fig.5. We shall see later that the integration of (4.9) is comparatively difficult in that region.

For Figs. 7 and 8 the step was 1. Again the method (4.2) produces a spurious limit cycle. It appears that the size of the spurious limit cycles obtained does not depend on the initial point but only on the step size.

The last example had

$$f_1 = y_2\,(2\, y_1^2 + y_2^2) \qquad (4.10)$$

$$f_2 = -y_1^3$$

and initial point (0,1). The results for $h = k = 0.5$ are depicted in Fig.9. The points corresponding to (4.1) are reasonably close to the true trajectory even when five orbits have been completed, while the method (4.2) once more yields an incorrect picture of the situation.

We conclude that for the problems considered the geometrically derived, circularly exact algorithm (4.1) is better suited than its standard counterpart.

5. VARIABLE STEP

In this Section we construct and test a variable-step version of the circularly exact method (4.1). It should be emphasized that our aim is to demonstrate the possibility of such a construction rather than to develop a sophisticated code.

We first derive a variable-step circularly exact predictor formula. Given $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$, $\underset{\sim}{F}_{n+1}$ and a positive number h_{n+1}, this formula will yield the point $\underset{\sim}{y}^p_{n+2}$ which satisfies $\|\underset{\sim}{y}^p_{n+2} - \underset{\sim}{y}_{n+1}\| = h_{n+1}$ and lies on the circle C_n determined by $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$, $\underset{\sim}{F}_{n+1}$. Fig.10 depicts the two-dimensional plane defined by the points $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$ and the vector $\underset{\sim}{F}_{n+1}$. If we denote by γ the angle between $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$ and $\underset{\sim}{F}_{n+1}$, then the central angle subtended in C_n by $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$ is 2γ. Therefore the angle between $\underset{\sim}{y}_n - \underset{\sim}{y}^p_{n+2}$ and $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}^p_{n+2}$ is γ. (Recall that an inscribed angle is equal to one half of the corresponding central angle.)

Next let δ be the angle between $\underset{\sim}{F}_{n+1}$ and $\underset{\sim}{y}^p_{n+2} - \underset{\sim}{y}_{n+1}$. Then the angle between $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$ and $\underset{\sim}{y}^p_{n+2} - \underset{\sim}{y}_{n+1}$ is $\delta + \gamma$, and consideration of the triangle with vertices $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$, $\underset{\sim}{y}^p_{n+2}$ leads to the conclusion that the angle between $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$ and $\underset{\sim}{y}^p_{n+2} - \underset{\sim}{y}_n$ is also δ. We have denoted by $\underset{\sim}{N}_{n+1}$ the unit normal vector to C_n at $\underset{\sim}{y}_{n+1}$.

We are now in a position to derive the required formula. We project $\underset{\sim}{y}^p_{n+2} - \underset{\sim}{y}_{n+1}$ onto $\underset{\sim}{F}_{n+1}$, $\underset{\sim}{N}_{n+1}$ to get

$$\underset{\sim}{y}^{p}_{n+2} - \underset{\sim}{y}_{n+1} = h_{n+1} \cos\delta\, \underset{\sim}{F}_{n+1} + h_{n+1} \sin\delta\, \underset{\sim}{N}_{n+1}. \tag{5.1}$$

The Gram-Schmidt procedure enables us to express the normal vector $\underset{\sim}{N}_{n+1}$ in terms of $\underset{\sim}{F}_{n+1}$, $\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n$ as follows:

$$\underset{\sim}{N}_{n+1} = \cot\gamma\, \underset{\sim}{F}_{n+1} - (h_n \sin\gamma)^{-1}(\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n), \tag{5.2}$$

where $h_n = \|\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n\|$. Next δ can be eliminated by use of the sine theorem in the triangle $\underset{\sim}{y}_n$, $\underset{\sim}{y}_{n+1}$, $\underset{\sim}{y}^{p}_{n+2}$:

$$h_n/\sin\gamma = h_{n+1}/\sin\delta. \tag{5.3}$$

Finally γ is related to $\underset{\sim}{F}_{n+1}$, $\underset{\sim}{y}_{n+1}$, $\underset{\sim}{y}_n$ by the formula

$$\underset{\sim}{F}^{T}_{n+1}\,(\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n) = h_n \cos\gamma. \tag{5.4}$$

When (5.2), (5.3), (5.4) are substituted into (5.1) the following predictor formula is obtained:

$$\underset{\sim}{y}^{p}_{n+2} = \underset{\sim}{y}_{n+1} + (h_{n+1}/h_n)^2[A_n\, \underset{\sim}{F}_{n+1} + \underset{\sim}{y}_n - \underset{\sim}{y}_{n+1}] \tag{5.5}$$

where

$$A_n = B_n + (B_n^2 - h_n^2 + \frac{h_n^4}{h_{n+1}^2})^{\frac{1}{2}}; \tag{5.6a}$$

$$B_n = \underset{\sim}{F}^{T}_{n+1}(\underset{\sim}{y}_{n+1} - \underset{\sim}{y}_n). \tag{5.6b}$$

Formula (5.5) reduces to formula (2.1) if $h_{n+1} = h_n$. It should also be noted that $\underset{\sim}{y}^{p}_{n+2}$ will not be defined if h_{n+1} is chosen larger than the diameter d_n of C_n. From Fig.10 this diameter is $h_n/\sin\gamma$, whence using (5.4), (5.6b) we obtain

$$d_n = h_n^2/(h_n^2 - B_n)^{\frac{1}{2}}. \tag{5.7}$$

In fact the algorithm we shall describe later imposes the condition $h_{n+1} \leq 0.5\, d_n$.

The corrector formula is written in the form

$$\underset{\sim}{y}_{n+2} = \underset{\sim}{y}_{n+1} + (h_{n+1}/\|\underset{\sim}{F}_{n+1} + \underset{\sim}{F}^{p}_{n+2}\|)(\underset{\sim}{F}_{n+1} + \underset{\sim}{F}^{p}_{n+2}), \qquad (5.8)$$

so that $\|\underset{\sim}{y}_{n+2} - \underset{\sim}{y}_{n+1}\| = h_{n+1}$.

In order to control the step-size a Milne device can be employed. Let $\underset{\sim}{y}^*$ be the point such that $\|\underset{\sim}{y}^* - \underset{\sim}{y}_{n+1}\| = h_{n+1}$ and $\underset{\sim}{y}^*$ lies on the trajectory through $\underset{\sim}{y}_{n+1}$. Then as in Section 3

$$\underset{\sim}{y}^* - \underset{\sim}{y}^{p}_{n+2} \simeq (1/6)(h^3_{n+1} + h^2_{n+1}\, h_n)[\dot{\kappa}\underset{\sim}{n} + \kappa\tau\underset{\sim}{b}], \qquad (5.9)$$

$$\underset{\sim}{y}^* - \underset{\sim}{y}_{n+2} \simeq (-1/12)\, h^3_{n+1}\, [\dot{\kappa}\underset{\sim}{n} + \kappa\tau\underset{\sim}{b}]\ , \qquad (5.10)$$

and elimination of the term in square brackets leads to

$$\underset{\sim}{y}^* - \underset{\sim}{y}_{n+2} \simeq [h_{n+1}/(3\, h_{n+1} + 2\, h_n)][\underset{\sim}{y}_{n+2} - \underset{\sim}{y}^{p}_{n+2}]. \qquad (5.11)$$

We considered the following algorithm

(1) Given $\underset{\sim}{y}_0, \underset{\sim}{y}_1, h_0, h_1$, $\varepsilon > 0$, with $h_0 = \|\underset{\sim}{y}_1 - \underset{\sim}{y}_0\|$ set $n = 0$;

(2) Evaluate $\underset{\sim}{F}_{n+1}$. Use (5,6b),(5.8) to compute B_n, $1/d_n$. If $1/h_{n+1} < 2/d_n$, set $h_{n+1} = d_n/2$;

(3) Compute $\underset{\sim}{y}^{p}_{n+2}$ according to (5.5), evaluate $\underset{\sim}{F}^{p}_{n+2}$ and form $\underset{\sim}{y}_{n+2}$ (formula (5.8)).

(4) Use (5.11) to estimate the error $e = \|\underset{\sim}{y}^* - \underset{\sim}{y}_{n+2}\|$. Set $h^*_{n+1} = h_{n+1}\,(e/\varepsilon)^{-\frac{1}{3}}$;

(5) If $e > \varepsilon$, set $h_{n+1} = h^*_{n+1}$ and go to (3);

(6) Print $\underset{\sim}{y}_{n+2}$, set $h_{n+2} = h^*_{n+1}$, $n = n+1$ and go to (2).

The trapezuidal rule in correction-to-convergence mode was used to compute $\underset{\sim}{y}_1$ and initialize the algorithm, which was tested with several tolerances ε and various initial points in the systems (4.9),(4.10).

The following three-dimensional system was also considered:

$$f_1 = y_2,$$
$$f_2 = y_1, \tag{5.12}$$
$$f_3 = 4\, y_1\, y_2.$$

Fig. 11 shows the results for the system (4.9) with $\varepsilon = 0.001$ and $\underset{\sim}{y}_0 = (30,30)$. The true trajectory starting from (30,20) is also depicted in order to display the rapid convergence of the integral curves in the vicinity of the vertical portion C,D. It is well-known that this convergence forces any explicit algorithm to take a small step. By comparison the step is larger along AB, where the neighbouring integral curves are almost parallel.

Fig.12 also refers to the system (4.9), but now $\varepsilon = 0.005$ and $\underset{\sim}{y}_0 = (0,1)$. The maximum Euclidean distance between consecutive points is 1.6.

Fig.13 corresponds to the system (4.10) with $\varepsilon = 0.001$ and $\underset{\sim}{y}_0 = (0,1)$.

The system (5.12) was integrated starting from (1,0,1). The true solution is given in parametric form by

$$y_1(t) = \cos t$$
$$y_2(t) = -\sin t \tag{5.13}$$
$$y_3(t) = \cos 2t$$

The integration was stopped when roughly a quarter of an orbit had been completed. This corresponds to an arclength of 2.63. The curvature is initially 2.0, decreasing to 0.1 and increasing again to 2.0. When the tolerance was 0.01, nine steps were taken and the final point lay at a distance of 0.025 from the true integral curve. When the tolerance was decreased to 0.0001, thirty-two steps were required and the final error was 0.003.

We wish to emphasize that the algorithm presented here can be easily adap-

ted to yield several geometrical elements of the trajectory such as tangent and normal vectors, curvature, arc length, etc.

ACKNOWLEDGEMENTS

The authors are grateful to have benefited from many discussions of this topic with Professor J D Lambert of the University of Dundee. Support from the following agencies and institutes is willingly acknowledged:

North Atlantic Treaty Organization, National Aeronautics and Space Administration (USA), British Council in Spain. This work was carried out at the University of Dundee (Scotland) and at the National Research Institute for Mathematical Sciences, Council for Scientific and Industrial Research, Pretoria, (South Africa).

/MR

REFERENCES

[1] ALLGOWER, E. and GEORG, K. Simplicial and continuation methods for approximating fixed points and solutions to systems of equations, SIAM Review 22, (1980), 28-85.

[2] LAMBERT, J.D. and McLEOD, R.J.Y. Numerical methods for phase-plane problems in ordinary differential equations. In: Numerical Analysis Proceedings, Dundee 1979. Springer, Berlin 1980.

[3] LAURIE, D.P. Equispacing numerical methods for trajectory problems. In: Proceedings of the Sixth South African Symposium in Numerical Analysis, Durban 1980. Computer Science Department, University of Natal, Durban 1980.

[4] KELLER, H.B. Constructive methods for bifurcation and nonlinear eigenvalue problems. In: Computing methods in applied sciences and engineering. Springer, Berlin 1979.

[5] RHEINBOLDT, W.C. Solution fields of nonlinear equations and continuation methods. SIAM J. Numer. Anal., 17(1980), 221-239.

[6] SHAMPINE, L.F. Limiting precision in differential equation solvers. II: Sources of trouble and starting a code. Math. Comp. 32 (1978), 1115-1122.

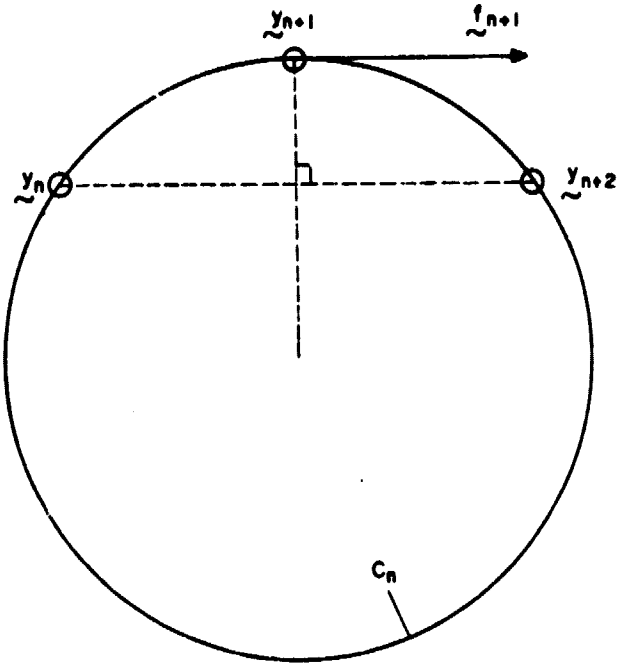
$\underset{\sim}{y}_{n+1}$
$\underset{\sim}{f}_{n+1}$
$\underset{\sim}{y}_n$
$\underset{\sim}{y}_{n+2}$
C_n

FIGURE 1

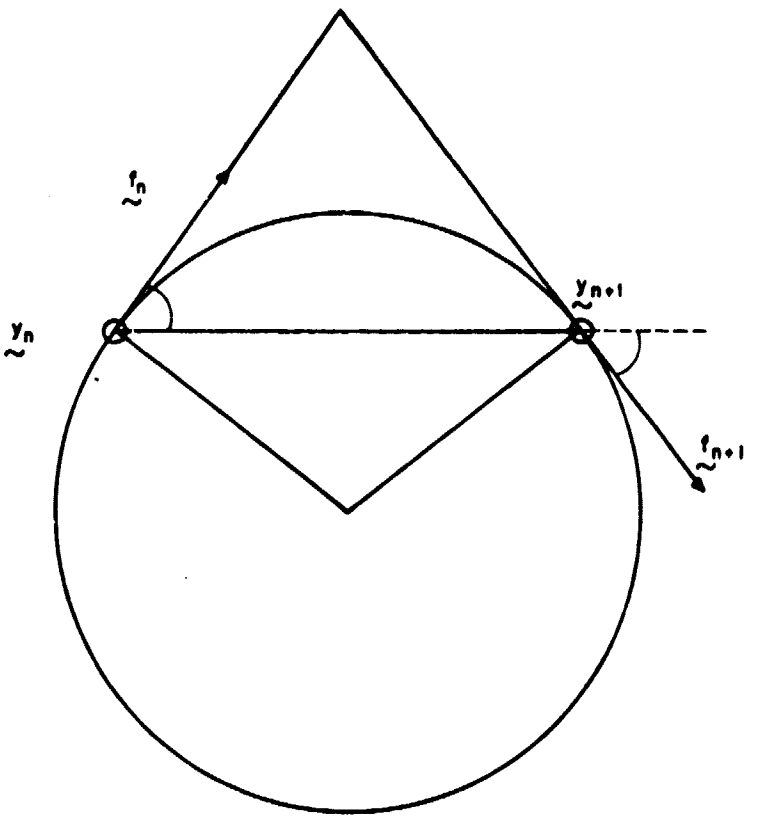
$\underset{\sim}{f}_n$
$\underset{\sim}{y}_{n+1}$
$\underset{\sim}{y}_n$
$\underset{\sim}{f}_{n+1}$

FIGURE 2

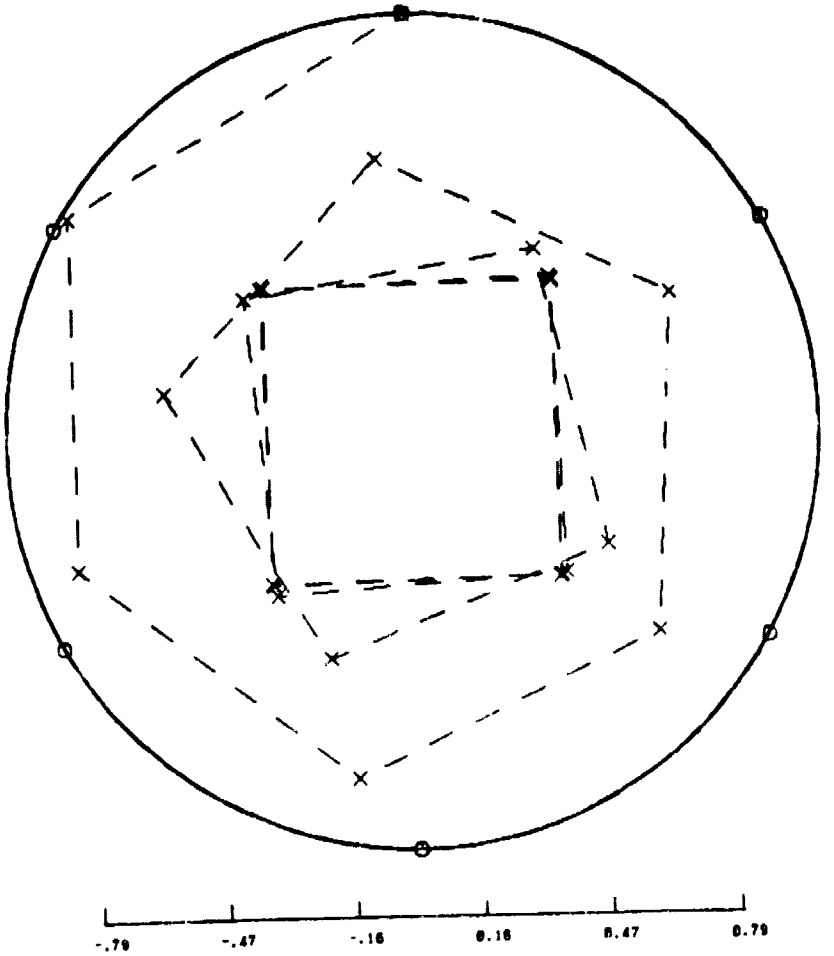

FIGURE 3

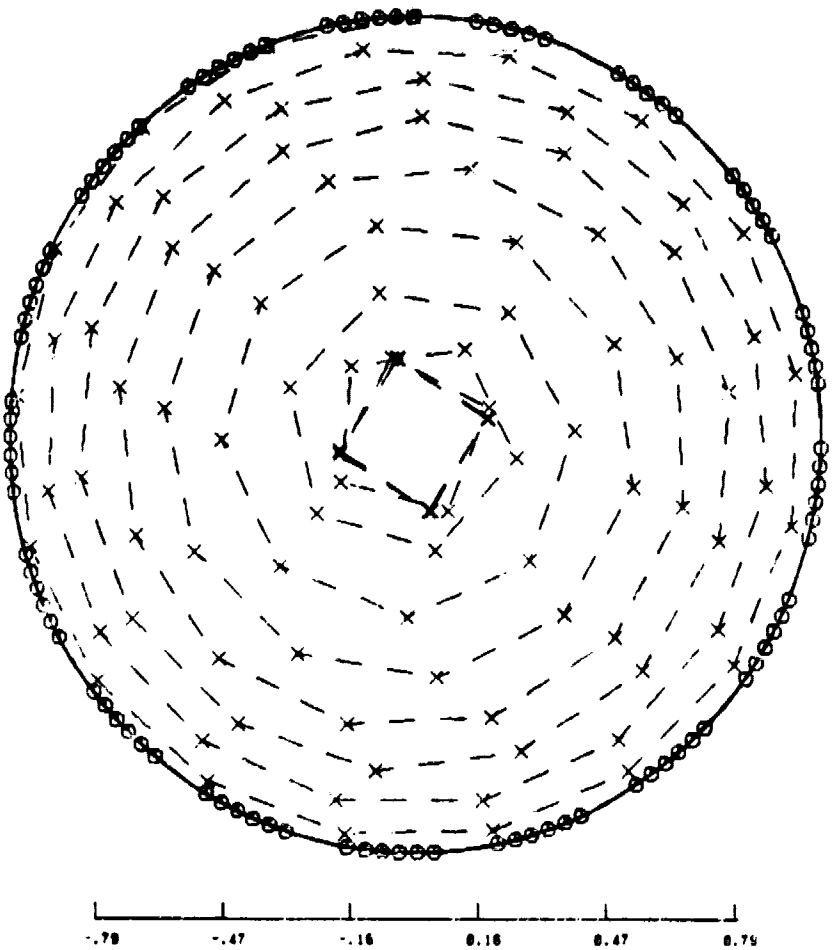

FIGURE 4

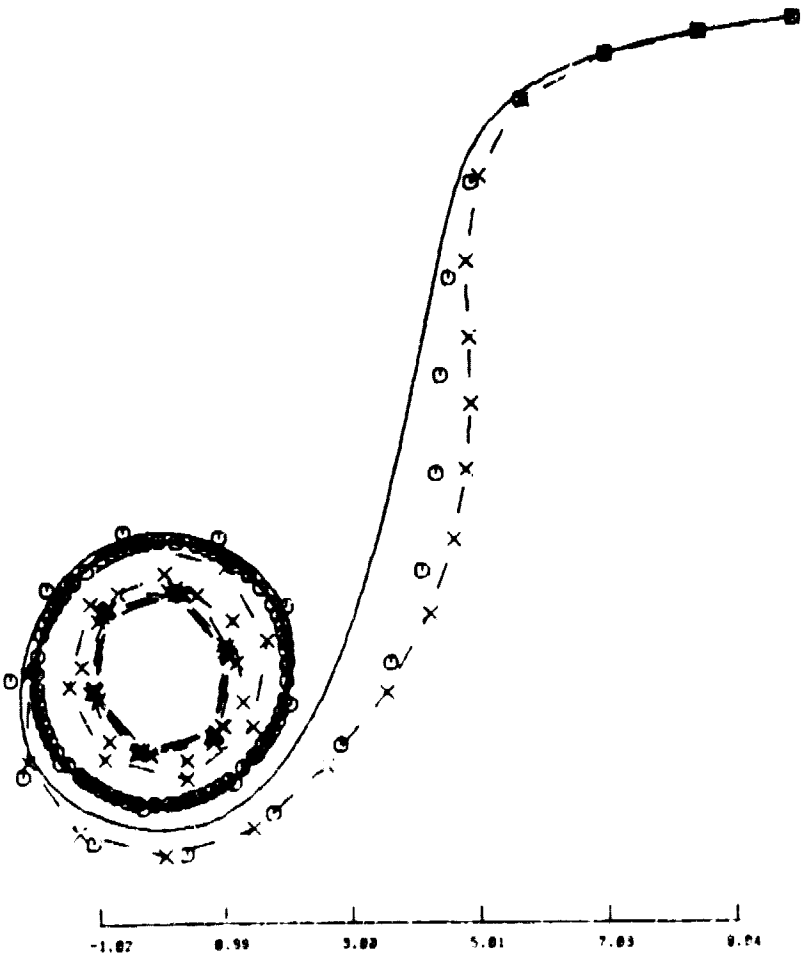

FIGURE 5

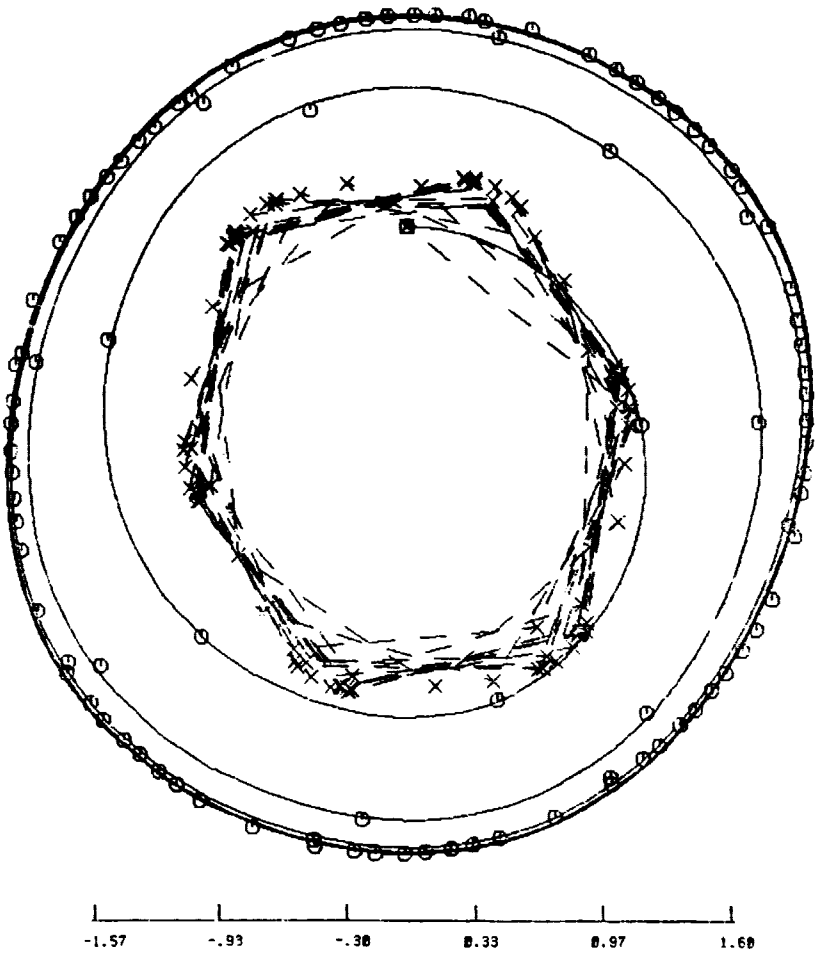

FIGURE 6

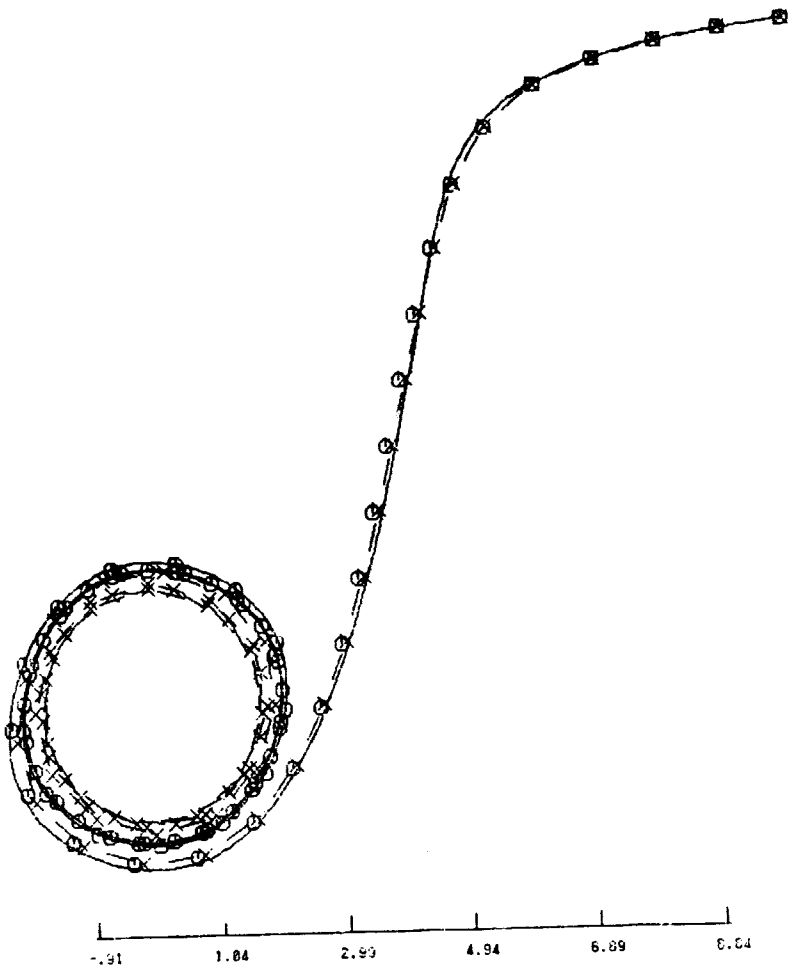

FIGURE 7

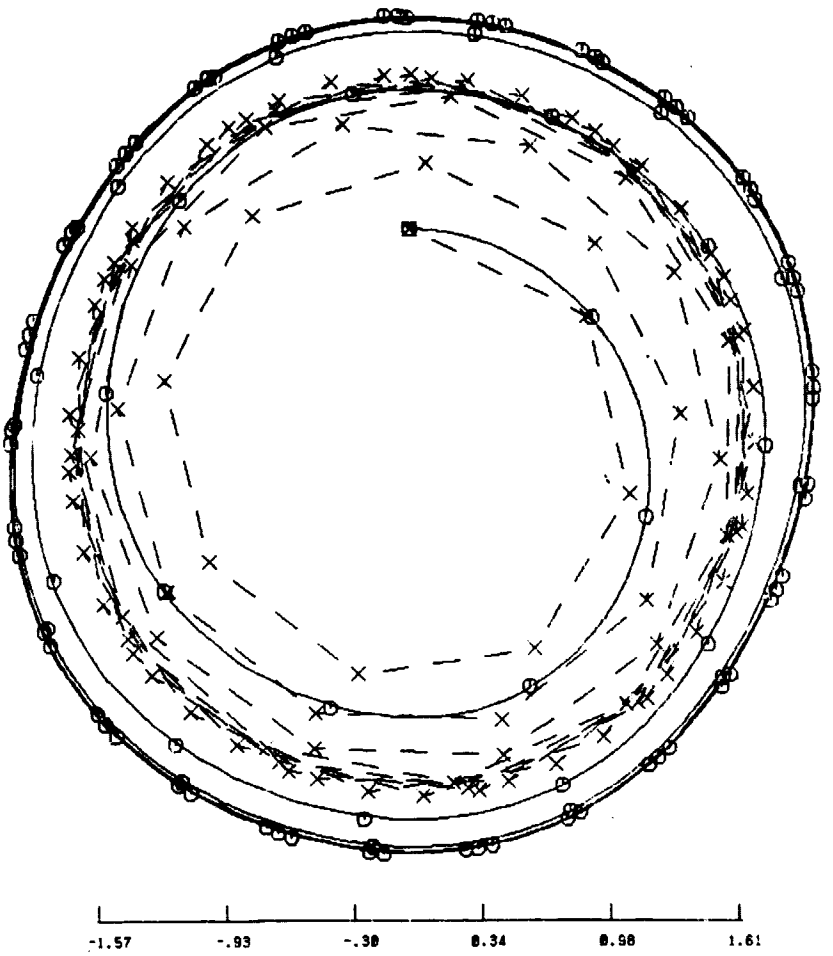

FIGURE 8

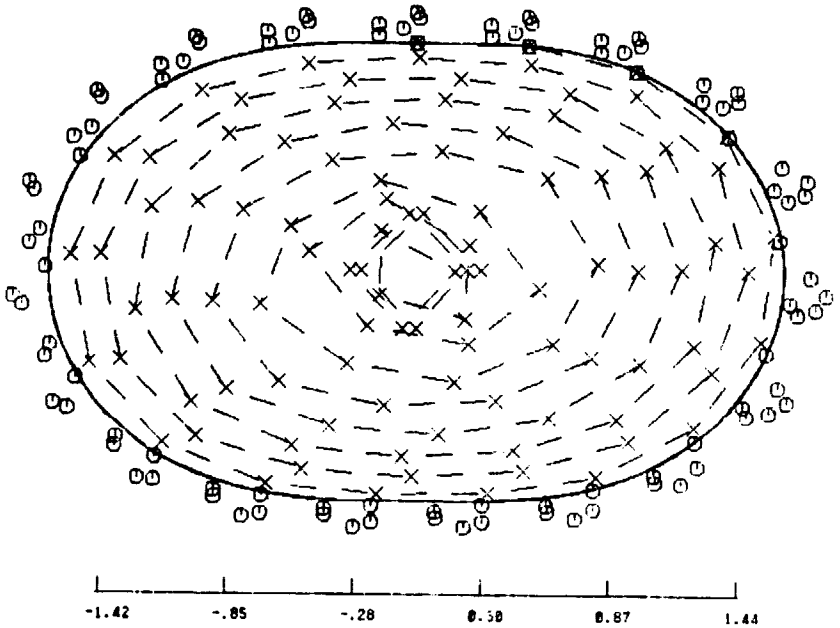

FIGURE 9

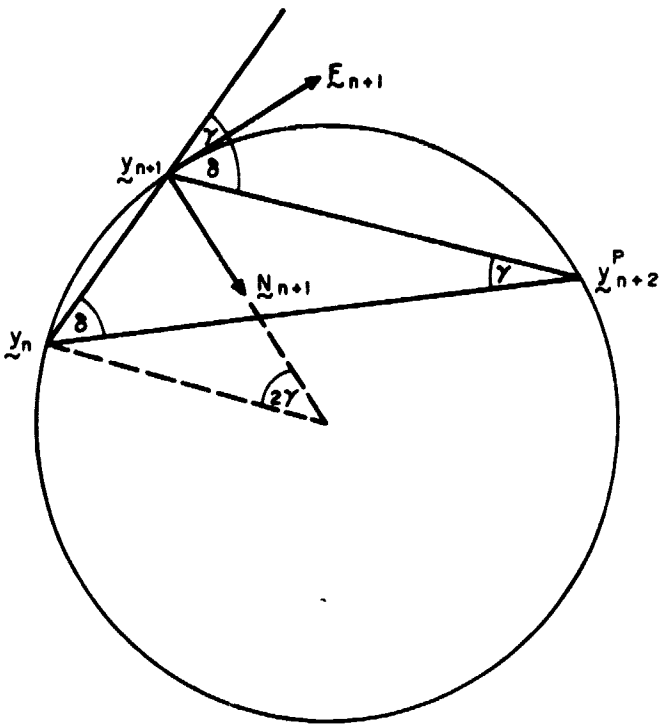
$\underset{\sim}{E}_{n+1}$
$\underset{\sim}{y}_{n+1}$
γ
δ
$\underset{\sim}{N}_{n+1}$
γ
$\underset{\sim}{y}^{P}_{n+2}$
δ
$\underset{\sim}{y}_{n}$
2γ

FIGURE 10

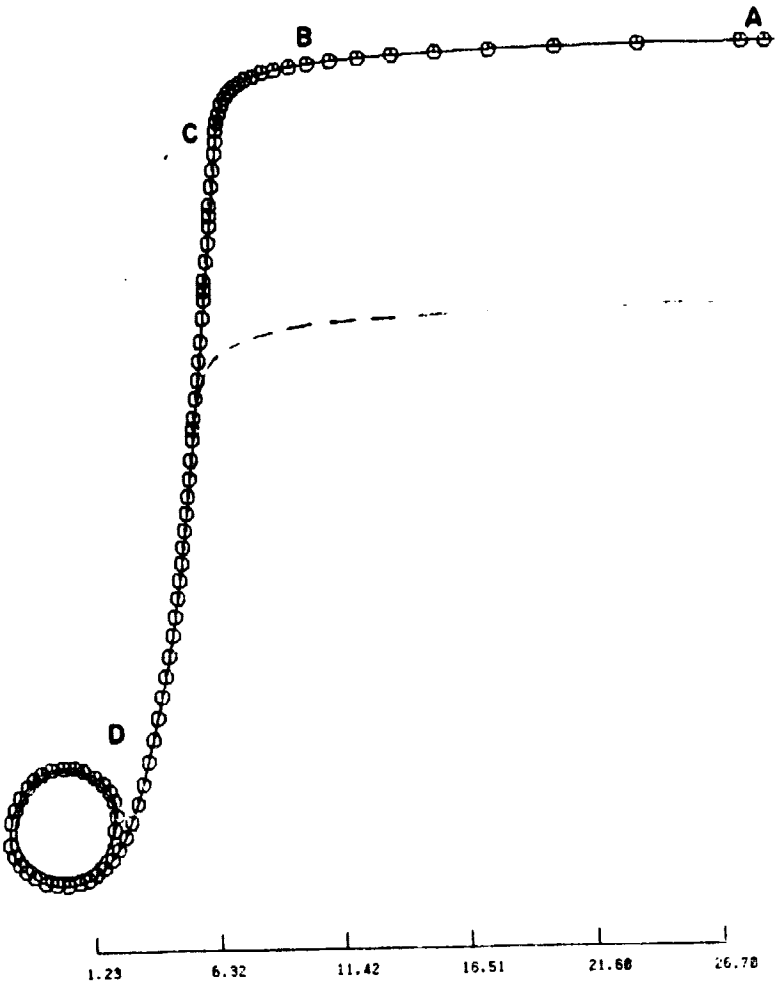

FIGURE 11

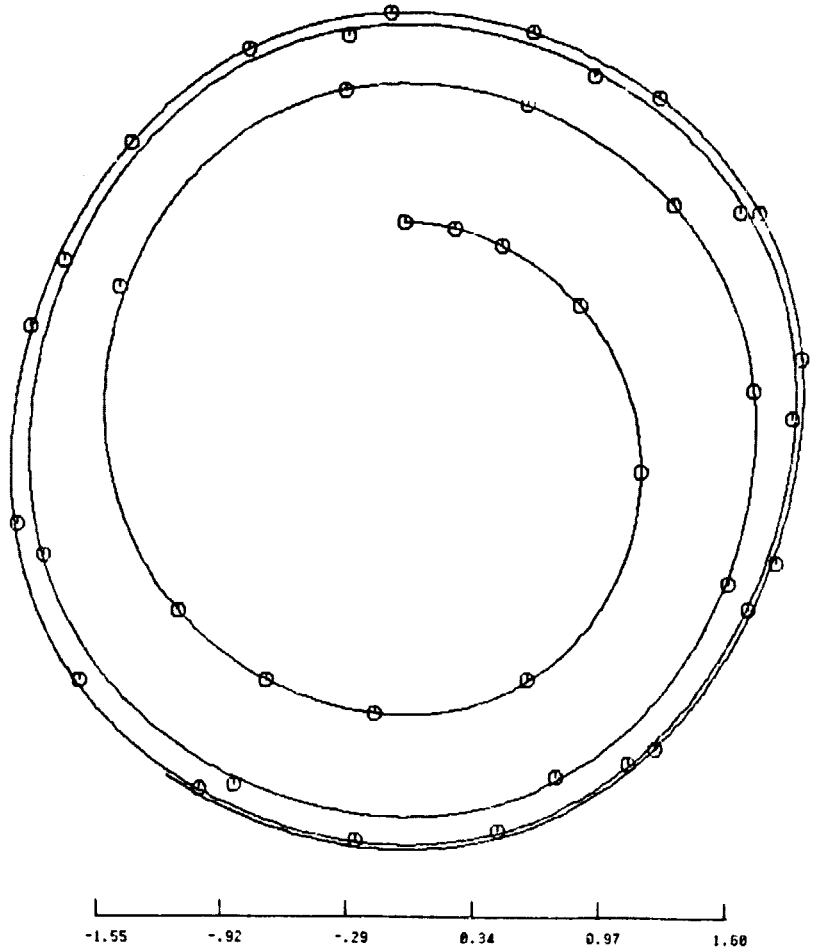

FIGURE 12

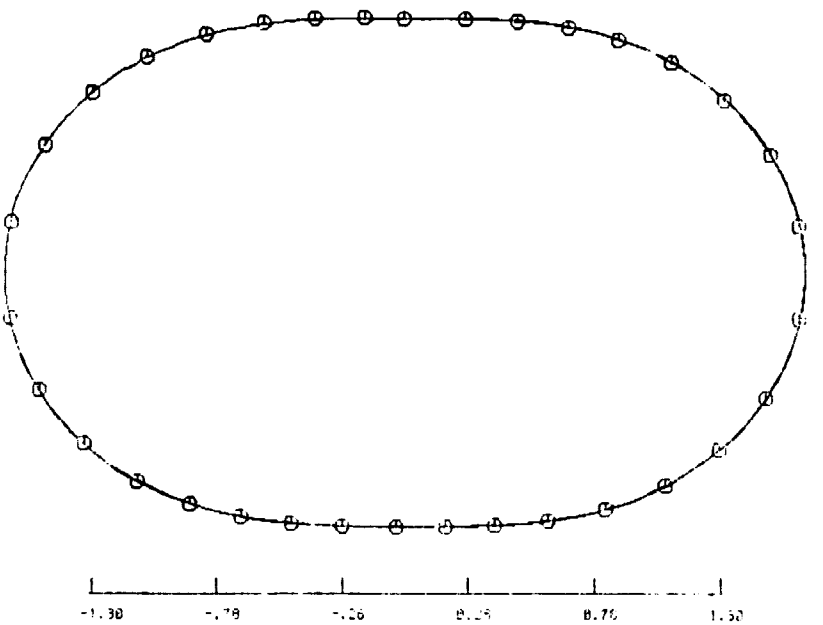

FIGURE 13