J. H. BARTLETT

MOTION UNDER A PERIODIC CUBIC FORCE

Det Kongelige Danske Videnskabernes Selskab
Matematisk-fysiske Meddelelser **36,** 11

Kommissionær: Munksgaard
København 1968

Synopsis

In order to ascertain the effects of non-linearities, a study has been made of the simple differential equation $\ddot{x} + p(t)\, x^3 = 0$, where $p(t)$ is a periodic square wave function of time. Integration was at first accomplished by the use of the Runge-Kutta method, but accuracy limited only by the computer capacity can be achieved by the construction of a table of elliptic functions.

When the time t is increased by one period, any point P of the $(x,\dot{x})$ phase plane is transformed into a point $T(P)$, defining a mapping of the phase plane into itself. The motion is rotatory when p is positive and hyperbolic when p is negative, so that iteration of T can result in wrinkling. As the distance from the origin increases, the motion becomes faster, and $T(P)$ becomes infinite. That is, a particle can go to infinity in a finite time.

The equation has an infinite number of periodic solutions, or the mapping has an infinite number of fixed points. For the $p(t)$ which we have chosen, there is at least one fixed point on the x-axis and on the $\dot{x}$-axis for each T^n $(n = 1, 2 \ldots)$, and a fixed point for T^{n+1} lies closer to the origin that one for T^n. These fixed points P, plus their transforms $T^k\, P$ $(k = 1 \ldots n-1)$ are apparently all that exist. A table is given of some fixed points with n small.

The inner part of the plane is rather stable, in the sense that particles have only a small chance of reaching infinity. If there is an invariant region around an elliptic fixed point for T^n, it may be found by obtaining the invariant curves issuing from the hyperbolic fixed points for T^n.

PRINTED IN DENMARK
BIANCO LUNOS BOGTRYKKERI A/S

Introduction*

Although non-linear differential equations have been studied extensively in connection with the dynamics of a particle, most of the literature has been devoted to systems where the Hamiltonian does not depend explicitly on time. However, the design of high-energy accelerators and the resolution of stability questions in celestial mechanics call for the development of adequate theory for Hamiltonians which are periodic in time. Basic existence theorems concerning invariant curves have been advanced recently by ARNOL'D (1) and by MOSER (2), but further work is apparently required before one has effective and practical methods of computation. In this paper, we shall discuss in detail the solution of a simple non-linear equation, with special reference to the location of the fixed points and to stability of solutions.

For one degree of freedom, the motion may be represented by the equation $\ddot{x}+p(t)\, f(x) = 0$, where $p(t)$ is periodic with period τ. For simplicity, we shall assume p to be a square wave, namely $p = p_0$ for $-\frac{\tau}{4} < t \leq \frac{\tau}{4}$ and $p = -p_0$ for $\frac{\tau}{4} < t \leq \frac{3\tau}{4}$, and $p_0 > 0$. If $f(x) = x$, the theory is well-known and given by FLOQUET (3). The equation with $f(x) = x + ax^3$ has been integrated numerically, for various initial conditions, by WRIGHT and POWELL (4) and discussed with the aid of perturbation theory (small values of "a") by MOSER (5). However, this work was very limited in extent, and a conversation with Dr. MOSER had as its subject the desirability of a theory for large values of x (or "a"). Subsequently, the author decided to neglect the linear term, to set $f(x) = x^3$, and to integrate with no reliance on perturbation theory. Such procedure brings out the effects due to the non-linearity of the system.

The Mathieu-Hill Equation

Before taking up the non-linear equation, it is necessary to review the linear case, where $f(x) = x$, because the literature is somewhat scattered. Our treatment will be based mainly on the work of COURANT and SNYDER (6).

* This work was reported in part to the USSR Mathematical Congress at Leningrad on July 5, 1961.

The Mathieu-Hill equation to be considered is

$$\frac{d^2x}{dt^2} + p(t)x = 0 \tag{1}$$

where $p(t)$ is defined as above.

Let the transformation T through one period be written

$$\left.\begin{aligned} x(\tau) &= M_{11}x(0) + M_{12}\dot{x}(0) \\ \dot{x}(\tau) &= M_{21}x(0) + M_{22}\dot{x}(0). \end{aligned}\right\} \tag{2}$$

This is a linear transformation of the phase plane $(x, \dot{x})$ into itself. It transforms the initial point $(1,0)$ into (M_{11}, M_{21}), and $(0,1)$ into (M_{12}, M_{22}). The coefficients M_{ij} must be obtained by actual integration of (1) for one period ($t = 0$ to $t = \tau$).

Since the Wronskian of the 2 solutions $(1,0)$ and $0,1)$ is constant and initially unity, the matrix M has determinant 1, and can be taken to be

$$M: \begin{pmatrix} \cos\sigma & \beta\sin\sigma \\ -(1/\beta)\sin\sigma & \cos\sigma \end{pmatrix} \tag{3}$$

since $p(t) = p(-t)$. If $|\cos\sigma| \leq 1$, then σ is real and Equations (2) may be written

$$\left.\begin{aligned} x(\tau) &= \quad x(0)\cos\sigma + \beta\dot{x}(0)\sin\sigma \\ \beta\dot{x}(\tau) &= -x(0)\sin\sigma + \beta\dot{x}(0)\cos\sigma. \end{aligned}\right\} \tag{4}$$

This shows that T just rotates the vector $(x, \beta\dot{x})$ through an angle σ, the magnitude of which is determined by $p(t)$ alone and not by the initial values of x and $\beta\dot{x}$. If we just consider the points $t = n\tau$, $n =$ integer, the quantity $x^2 + \beta^2\dot{x}^2$ is constant. When σ is real, β is real, the locus of the vector $(x, \dot{x})$, for $t = n\tau$, is an ellipse, and the origin is called an *elliptic fixed point*. If $n\sigma = 2\pi m$, with m and n integers, the initial point P and the companion points TP, T^2P, etc. will be fixed under T^n. If $n\sigma \neq 2\pi m$, repeated application of T generates more and more points on the ellipse.

When $|\cos\sigma| > 1$, σ is imaginary, and β must be imaginary also, since M_{12} is real. Put $\sigma = i\psi$ and $i\beta = \gamma$, a real quantity. Equation (4) becomes

$$\left.\begin{aligned} x(\tau) &= x(0)\,ch\,\psi + \gamma\dot{x}(0)\,sh\,\psi \\ \gamma\dot{x}(\tau) &= x(0)\,sh\,\psi + \gamma\dot{x}(0)\,ch\,\psi. \end{aligned}\right\} \tag{5}$$

For the points $t = n\tau$, the quantity $x^2 - \gamma^2\dot{x}^2$ is constant, and the locus consists of the branches of a hyperbola. For this case, the origin is called a *hyperbolic* fixed point.

The asymptotes to the hyperbolas are given by $x = \gamma\dot{x}$ and $x = -\gamma\dot{x}$. If the initial point P is on the first asymptote, then the map TP will also be, but the distance from the origin will have been stretched by the factor $\lambda_1 = ch\psi + sh\psi$, as we see by substitution into (5). Similarly, the distance from the origin to a point on the second asymptote will be shrunk, the factor now being $\lambda_2 = ch\psi - sh\psi$.

If $\dot{x}(0) = 0$, then Equation (4) yields $\cos\sigma = x(\tau)/x(0)$. Consequently, if $x \neq 0$, $\dot{x} = 0$ is a fixed point, we can test whether it is elliptic or hyperbolic by seeing how a point on the x-axis maps under T. If $|x(\tau)| < |x(0)|$, the fixed point is elliptic; if $|x(\tau)| > |x(0)|$, the fixed point is hyperbolic.

The x-axis is mapped into $x(\tau) = x(0)\, ch\psi$, $\gamma\dot{x}(\tau) = x(0)\, sh\psi$, the first equation giving $ch\psi$, and the second γ, assuming that the mapping has been done numerically. The stretch factors λ_1 and λ_2 follow immediately from $ch\psi$.

The Cubic Equation

Let us now consider the equation

$$\frac{d^2x}{dt^2} + p(t)x^3 = 0 \tag{6}$$

where $p(t)$ has period τ and

$$p = p_0 \quad \text{for} \quad -\frac{\tau}{4} < t \leq \frac{\tau}{4}, p_0 > 0,$$

$$p = -p_0 \quad \text{for} \quad \frac{\tau}{4} < t \leq \frac{3}{4}.$$

As in the case of the Mathieu-Hill equation, the important thing is to find the transformation T, or how the $(x,\dot{x})$ plane maps into itself when t is increased by the period τ. This mapping depends only on the initial values of x and $\dot{x}$, and not on t. For the Mathieu-Hill equation, conclusions about stability are immediate, and depend only on the value of σ, so that an elliptic point is stable and a hyperbolic point unstable. However, for the cubic equation and other similar non-linear equations, the mapping is not simple and conclusions about stability are more difficult to reach. Hyperbolic points are obviously unstable, while ARNOL'D and MOSER (loc. cit.) have presented proofs that elliptic points are stable if certain criteria are satisfied. For any given differential equation, one must examine the mapping to find out whether these conditions are in fact met.

If extremely high accuracy is not necessary, integration by the Runge-Kutta method is simple and expeditious. Our first results were obtained in this way, but these have been later improved by the use of Jacobian elliptic functions, in a manner to be discussed presently.

A table of these functions is stored on a computer, and the integration reduces to looking up these functions and their inverses in the table. The accuracy is then limited by the number of bits which can be stored in a computer location.

First Integration

When $p = p_0$, the first integral of Eq. (6) is

$$\dot{x}^2 + (p_0/2)x^4 = (p_0/2){x_0}^4 \tag{7a}$$

where x_0 is the intercept on the x-axis.

When $p = -p_0$, and there is an intercept on the x-axis at x_1, the first integral of Eq. (6) is

$$\dot{x}^2 - (p_0/2)x^4 = -(p_0/2){x_1}^4. \tag{7b}$$

When $p = -p_0$, and there is an intercept on the $\dot{x}$-axis, the first integral is

$$\dot{x}^2 - (p_0/2)x^4 = (p_0/8)a^4 \tag{7c}$$

where a is a real constant.

The transition from (7b) to (7c) occurs when

$$\dot{x} = \pm (p_0/2)^{1/2}x^2. \tag{8}$$

If this condition is satisfied initially, the motion will be along a parabola passing through the origin. Otherwise, the motion is along a quasi-ellipse (7a), or a quasi-hyperbola (7b or 7c).

Reduction to Standard Form

In Equation (7a), let $x = zx_0$. Then

$$\dot{z}^2 = (p_0/2){x_0}^2(1 - z^4). \tag{9a}$$

In Equation (7b), let $x = x_1/z$. Then

$$\dot{z}^2 = (p_0/2)x_1{}^2(1 - z^4). \tag{9b}$$

Finally in Equation (7c), let $x^2 = a^2w^2/2$ and $w^2 = (1-z)/(1+z)$. Since $w\dot{w} = -\dot{z}/(1+z)^2$, we obtain

$$\dot{z}^2 = (p_0/2)a^2(1 - z^4). \tag{9c}$$

These equations all have the same form, and differ only in the parameter which enters.

Jacobian Elliptic Functions

Since the integration of Equations (9) involves Jacobian elliptic functions with parameter $k = (1/2)^{1/2}$, we enumerate for future reference some of the properties of these functions.

$$sn\,u = u - \frac{1}{4}u^3 + \frac{11}{160}u^5 \ldots \tag{10a}$$

$$cn\,u = 1 - \frac{1}{2}u^2 + \frac{1}{8}u^4 - \frac{3}{80}u^6 \ldots \tag{10b}$$

$$dn\,u = 1 - \frac{1}{4}u^2 + \frac{3}{32}u^4 \ \ldots \tag{10c}$$

$$d\,sn\,u = cn\,u\,dn\,u\,du \tag{10d}$$

$$d\,cn\,u = -\,sn\,u\,dn\,u\,du \tag{10e}$$

$$2d\,dn\,u = -\,sn\,u\,cn\,u\,du \tag{10f}$$

$$cn^2u + sn^2u = 1 \tag{10g}$$

$$dn^2u + \frac{1}{2}sn^2u = 1 \tag{10h}$$

$$d^2\,cn\,u/du^2 = -\,cn^3u \tag{10i}$$

$$\frac{d^2}{du^2}\frac{sn\,u}{1 - cn\,u} = \frac{1}{2}\left(\frac{sn\,u}{1 - cn\,u}\right)^3. \tag{10j}$$

The functions $sn\,u$ and $cn\,u$ may be regarded as distorted sines and cosines, with period $4K = 7.41629871$. The function $dn\,u$ is even, and also symmetric about $u = K$, with $dn^2K = \frac{1}{2}$. Note that $cn(2K-u) = -\,cn\,u$. Also

$$D\,sn(a+b) = sn\,a\,cn\,b\,dn\,b + sn\,b\,cn\,a\,dn\,a \tag{10k}$$

$$D\,cn(a+b) = cn\,a\,cn\,b - sn\,a\,sn\,b\,dn\,a\,dn\,b \tag{10l}$$

$$D\,dn(a+b) = dn\,a\,dn\,b - \frac{1}{2}sn\,a\,sn\,b\,cn\,a\,cn\,b \tag{10m}$$

where

$$D = 1 - \frac{1}{2}sn^2a\,sn^2b. \tag{10n}$$

Furthermore,

and

$$\left.\begin{aligned} sn\,iv\,cn\,v &= i\,sn\,v, \quad cn\,iv\,cn\,v = 1, \\ dn\,iv\,cn\,v &= dn\,v. \end{aligned}\right\} \tag{10p}$$

Second Integration

The integral of Eq. (9a) is

$$z = cn(p_0^{1/2}x_0t) \tag{11}$$

as may be verified by using (10e), (10g), and (10k). Thus, if the motion starts on the x-axis at x_0 and is of duration t, the final point (quasi-elliptic motion) will be at

with

$$\left.\begin{aligned} x &= x_0cn(p_0^{1/2}x_0t) = x_0cn\,u \\ \dot{x} &= -p_0^{1/2}x_0^2sn\,u\,dn\,u \\ u &= p_0^{1/2}x_0t. \end{aligned}\right\} \tag{12}$$

If the motion starts on the x-axis at x_1 and is quasi-hyperbolic, the final point will be at

with

$$\left.\begin{aligned} x &= x_1/cn\,u \\ \dot{x} &= p_0^{1/2}x_1^2sn\,u\,dn\,u/cn^2\,u \\ u &= p_0^{1/2}x_1t. \end{aligned}\right\} \tag{13}$$

Finally, if quasi-hyperbolic motion starts on the $+\dot{x}$-axis, the point at time t will be at

with

$$\left.\begin{aligned} x &= (a^2/2)^{1/2}(1 - cn\,u)/sn\,u \\ &= (a^2/2)^{1/2}sn\,u/(1 + cn\,u) \\ \dot{x} &= (p_0/2)^{1/2}a^2\,dn\,u/(1 + cn\,u) \\ u &= p_0^{1/2}a\,t. \end{aligned}\right\} \tag{14}$$

If the motion starts from a point not on either axis, the appropriate value of the parameter (x_0, x_1, or a) must first be found from Equation (7), and the time from the axis to the initial point can be calculated by solving Eq. (12), (13), or (14). This gives immediately the time from the axis to the final point, since the duration of the motion is presumed to be given. The coordinates of the final point can now be calculated directly from Eq. (12), (13), or (14).

In what follows, it will be convenient to denote quasi-elliptic mapping for time $\tau/4$ by e, and quasi-hyperbolic mapping for this time by h. Also, from now on, we shall omit the term "quasi".

Scale

The values of p_0 and of τ can be fixed quite arbitrarily. Inspection of (12) shows that the motion depends essentially on $u = p_0{}^{1/2} x_0 t$, so that if we select p_0 and t, then x_0 may be calculated from u. Powell and Wright (4) chose $p_0 = e/3$, where $e = 0.1$, and the quarter period $\tau/4 = 1.5$. In our calculations, it has been convenient to modify this slightly, taking $p_0 = 0.037$ and $\tau/4 = 1.5$.

General Properties of the Motion

Symmetry and Reversal

If $x(t)$ is a solution of Equation (6), so are $x(-t)$, $-x(t)$, and $-x(-t)$. The motion may be reversed by replacing $(x, \dot{x})$ by $(x, -\dot{x})$. An equivalent reverse motion is $(-x, \dot{x})$. To reverse the motion, reflect in the x-axis or the $\dot{x}$-axis.

Elliptic Motion

If the motion is elliptic for a time $\tau/4$, the x-axis maps into the curve $e(x)$. This starts as a cubic from the origin, since $x = x_0$ and $\dot{x} \simeq -p_0 x_0{}^3 \tau/4$, and then turns into a curve which spirals outward and clockwise, traversing one quadrant each time that $p_0{}^{1/2} x_0 \tau/4$ increases by K. The first intercept on the $\dot{x}$-axis will be, by (12), at $\dot{x} = -(p_0/2)^{1/2} x_0{}^2$, where $x_0 = 4K/p_0{}^{1/2} = 6.4259$, since $u = K$ for $x = 0$.

Under elliptic motion, the time to traverse one quadrant from the x-axis is $K/x_0 p_0{}^{1/2}$, and therefore *the outer parts of the plane are rotated faster than the inner parts.*

If x_0 is fixed, the points reached by elliptic motion lie on the oval described by (7a) and (12), and only differ by the value of the parameter u, which will be called the *phase*. Under elliptic motion, the phase difference between 2 points on the above oval remains constant. Since the $\dot{x}$-axis is described by $u = -K$, the map $e(\dot{x})$ will also be a curve which spirals outward and clockwise, each point on it lagging in phase by the amount K behind a corresponding point on $e(x)$. For small values of $\dot{x}$, the locus $e(\dot{x})$ is given by $x = \dot{x}_0\tau/4$, or a straight line.

Hyperbolic Motion

If the motion is hyperbolic for a time $\tau/4$, the x-axis maps into the curve $h(x)$, which starts out from the origin as a cubic, namely $x = x_0$ and $\dot{x} \simeq p_0 x^3 \tau/4$. The maximum value of u equals K, and if $v = K - u$ is small, Equation (13) yields

$$\dot{x} \simeq (p_0/2)^{1/2} x^2 \tag{15}$$

and so the cubic approaches this parabola asymptotically from below as $u \to K$.

The corresponding map of the $\dot{x}$-axis, $h(\dot{x})$, starts out from the origin as a straight line $x = \dot{x}_0\tau/4$. Equation (14) shows that x becomes infinite when $u = 2K$ and also that $x^2 = (a^2/2)\,(1 - cn\,u)/(1 + cn\,u)$ and

$$\dot{x} = 2(p_0/2)^{1/2} x^2 dn\,u/(1 - cn\,u).$$

But, as $u \to 2K$, this last equation becomes

$$\dot{x} \to (p_0/2)^{1/2} x^2$$

so that the straight line changes into a curve $h(\dot{x})$ which also approaches the parabola (15) asymptotically, but from above.

Under hyperbolic motion, a finite point (not the origin) will move to infinity in a finite time t_0. If the point starts at x_1 on the x-axis, then from (13) we have $t_0 = K/p_0^{1/2} x_1$. If the point starts from the $\dot{x}$-axis, then from (14) $t_0 = 2K/p_0^{1/2} a$, where $a^2 (p_0/2)^{1/2} = 2\dot{x}_0$. It should then be possible to fix a value for t_0, and to find a curve C in the $(x,\dot{x})$ plane such that any point on this curve will just reach infinity in the time t_0.

Part of curve C may be regarded as generated by points leaving the x-axis and travelling for time $(K/p_0^{1/2} x_1) - t_0$. If $v = p_0^{1/2} x_1 t_0$, then

$$\left.\begin{aligned} x &= x_1/cn(K-v) = \sqrt{2}\,x_1 dn\,v/sn\,v \\ &= (2/p_0)^{1/2} v\,dn\,v/t_0 sn\,v \\ \dot{x} &= (2/p_0)^{1/2} v^2\,cn\,v/t_0{}^2 sn^2\,v. \end{aligned}\right\} \tag{16}$$

and

(The formula for $\dot{x}$ is obtained from Equation (13) by noting that $sn(K-v) = cn\,v/dn\,v$ and $dn(K-v) = \sqrt{1/2}\,dn\,v$).

Now equation (7c) will reduce formally to (7b) if we put $a = (1+i)x_1$. The remainder of curve C, that generated by points leaving the $\dot{x}$-axis, can then be found by setting $u = (1+i)v$ and finding what Equations (16) become in terms of u. Making use of Equations (10k)–(10p) and the relation

$$(2dn^2v - isn^2v)/(1+icn^2v) = 1-i,$$

we find

$$\left.\begin{aligned} \frac{2v\,dn\,v}{sn\,v} &= \frac{u\,(1+cn\,u)}{sn\,u} = \frac{u\,sn\,u}{1-cn\,u}, \\ \frac{v\,cn\,v}{sn\,v\,dn\,v} &= \frac{u\,dn\,u}{sn\,u}, \end{aligned}\right\} \tag{17a}$$

and

$$\frac{2v^2\,cn\,v}{sn^2\,v} = \frac{u^2\,dn\,u}{1-cn\,u}. \tag{17b}$$

Consequently, substituting (17a) and 17b) in (16), we have

$$\left.\begin{aligned} x &= (2/p_0{}^{1/2})\,u\,sn\,u/2t_0(1-cn\,u) \\ \dot{x} &= (2/p_0{}^{1/2})\,u^2 dn\,u/2t_0(1-cn\,u) \\ u &= p_0{}^{1/2} a\,t_0. \end{aligned}\right\} \tag{18}$$

with

(These formulas can also be obtained by using $2K-u$ instead of u in the arguments of Equations (14)).

Equations (16) and (18) represent together a smooth curve. It has a companion, obtained by changing x to $-x$ and $\dot{x}$ to $-\dot{x}$. Any point outside of these two curves will, under hyperbolic motion, reach infinity in time less than t_0. For the present problem, the maximum time that hyperbolic motion takes place continuously is one half-period, so that we shall set $t_0 = \tau/2 = 3$. All points which can reach infinity in time less than this will be called primarily unstable, and the above curves mark the boundary of primary instability. They are shown in Figure 1. The x—intercept has

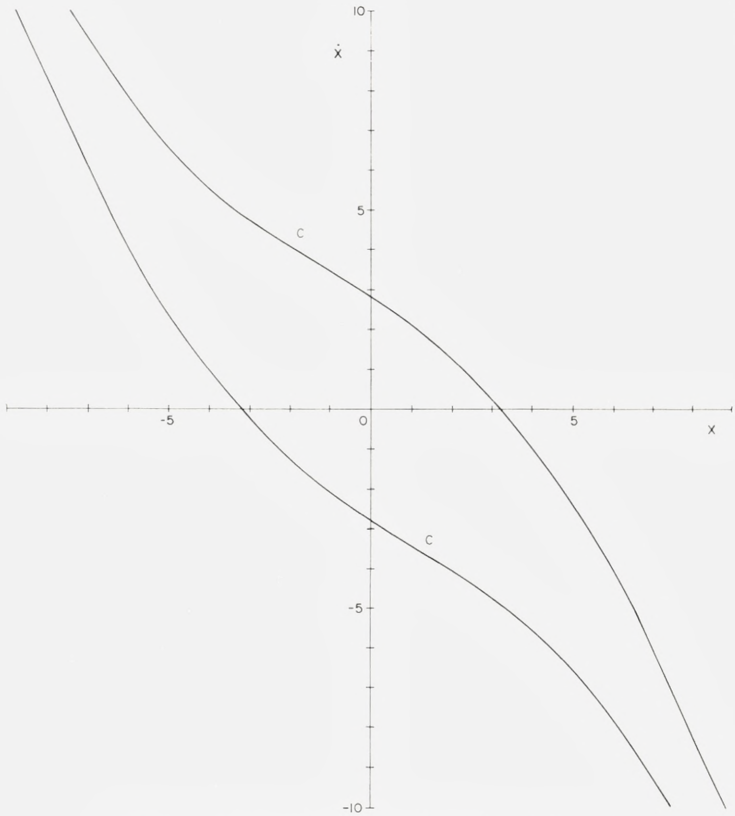

Figure 1. Primary Instability Boundary C, ($\dot{x}$ vs x).

magnitude $K/3p_0^{1/2} = 3.21296$, while the $\dot{x}$ – intercept has magnitude $(2/p_0)^{1/2}(K/t_0)^2 = 2.8082$.

The behavior of Equation (16) or (18) as $x \to \infty$ is readily obtained. From (16),

$$\dot{x} = (p_0/2)^{1/2} x^2 \, cn \, v / dn^2 \, v.$$

Letting $u \to 2K$, we have $x \to \infty$ and

$$\dot{x} = -(p_0/2)^{1/2} x^2. \tag{15a}$$

In the limit, curves C approach either this parabola or that described by (15) with $x < 0$. (When there is no ambiguity, we shall just say "the parabola").

The boundary between the two parts of curve C is marked by the 2 points obtained by setting $x_1 = 0$, or $v = 0$, or $u = 0$. Equation (16) then gives $x = (2/p_0)^{1/2}/t_0$ and $\dot{x} = (p_0/2)^{1/2}x^2$. The motion, as a function of t_0, is along the parabola, inward if $\dot{x}/x < 0$ and outward if $\dot{x}/x > 0$. The velocity is proportioned to x^2 and is thus faster the farther out we are. Knowing the initial x, we can calculate the initial t_0, add an increment Δt_0, and then obtain the final value of x.

Hyperbolic motion, which has this parabolic motion as a special case, *is thus faster the farther out in the plane the points are.*

The combined motion, which is always such as to preserve area in the phase plane, is rotatory half the time and hyperbolic for the other half. The inward hyperbolic motion ($\dot{x}/x < 0$) compresses toward the origin, and the outward hyperbolic motion ($\dot{x}/x > 0$) pulls away from the origin, in a rough manner of speaking. Wrinkling is the result.

Combined Motion (small x_0)

Let the motion start at a small value of x_0 on the x-axis, and be elliptic for the time $t_0 = \tau/4$, hyperbolic for time $\tau/2$ and elliptic for time $\tau/4$. We wish to find the locus of the end point $e\ h^2\ e\ (x_0) = T_e(x_0)$. Let x_a = abscissa after time $\tau/4$ from x_0, and t_1 = time from x_a along the hyperbolic part to the axis at x_1. The values of x will be taken so small as to be approximately constant. From Equations (12) and (13),

$$\left.\begin{aligned} x_a &= x_0\, cn(p_0{}^{1/2} x_0 t_0) = x_1/cn(p_0{}^{1/2} x_1 t_1) \\ &\simeq x_0 - \frac{1}{2} p_0 x_0{}^3 t_0{}^2 \simeq x_1 + \frac{1}{2} p_0 x_1{}^3 t_1{}^2 \end{aligned}\right\} \tag{19}$$

and

$$\dot{x}_a \simeq -p_0 x_0{}^3 t_0 \simeq -p_0 x_1{}^3 t_1. \tag{20}$$

If $\lambda = p_0 x_0{}^2$, then, from Equation (19)

$$x_1 \simeq x_0 - \lambda x_0 t_0{}^2 = x_0(1 - \lambda t_0{}^2)$$

and from Equation (20)

$$(t_1/t_0) = (x_0/x_1)^3 \simeq 1 + 3\lambda t_0{}^2$$

or

$$t_1 - t_0 = 3\lambda t_0{}^3 = 3\, p_0 x_0{}^2 t_0{}^3.$$

This is the time required for the hyperbolic motion from $he(x_0)$ to x_1. The time for the (elliptic) motion to go from the x-axis to $T_e(x_0)$ is very closely twice this. Substituting in Equation (20), the final value of $\dot{x}$ will be $\dot{x}_f = -p_0 x_0^3 (6 p_0 x_0^2)(\tau/4)^3 = -6(\tau/4)^3 p_0^2 x_0^5$.

The mapping $e\, h^2\, e$ thus, for small x_0, displaces a point on the x-axis downwards by an amount proportional to x_0^5. Calling this mapping temporarily $T(x_0)$, it is obvious that $T^2(x_0)$ will, for small positive x_0, lie below $T(x_0)$, and so on. This mapping is then a clockwise twist of the plane, the twist increasing with distance from the origin.

Fixed Points

A transformation T will in general carry one point P of the plane into another point $T\,(P)$, but there do exist special points, called *fixed* points, such that $T\,(P) = P$. The corresponding solutions of the differential equation are *periodic*, since T refers to the change due to a replacement of t by $t+\tau$. (The motion is to start in the middle of a sector, with p positive or negative). Let the symbol (n/m) signify that a point is fixed under T^n, and that the origin has been encircled exactly m times. As will be shown below, the points $(1/1)$, which are fixed under T, lie either on the x-axis or the $\dot{x}$-axis.

The Transformation $T = e\, h^2\, e(x_0)$

As x_0 increases, the curve $e(x_0)$ spirals outward and clockwise, intersecting the parabola when $cn^4\, u = 1/2$ i. e. when $x_0 = 2.037$. At this stage $he(x_0)$, which starts out from the origin as a fifth-power downwards, has turned and reached the parabola. Further increase of x_0 puts $e(x_0)$ across the parabola, and with it $he(x_0)$. The curve $e(x_0)$ reaches the $\dot{x}$-axis when $x_0 = 6.426$. On the other hand, $h^{-1}(\dot{x})$, for $\dot{x} < 0$, starts out as a straight line downward and to the right, eventually approaching the parabola asymptotically, so that it will intersect $e(x_0)$, and does so for $x_0 = 3.414$ (calculated). At this point, the motion is symmetric with respect to the $\dot{x}$-axis, so that $e\, h^2\, e(x_0) = -x_0$, which means that this value of x_0 is fixed under T^2. We also see that $T(x_0)$ has a spiral nature and has reached the negative x-axis. The next important event that occurs is that $e(x_0)$ intersects the primary instability boundary (P. I. B.) (see Figure 1), so that $h^2\, e(x_0)$ becomes infinite. What happens is that $h^2\, e(x_0)$ approaches the parabola in the second quadrant, with the magnitude of x increasing as x_0 does. The subsequent elliptic motion becomes faster and faster, with $e\, h^2\, e(x_0)$ intersecting the $+\,x$-axis and spiralling to infinity. If the value of x for which this intersection first takes place is x_1, then $T(x_0) = x_1$, and the points x_0 and x_1 are fixed under T^2, and classified as (2/2). We thus obtain a whole series of fixed points as x_0 approaches the limiting value where $e(x_0)$ intersects the P.I.B. Later, however, as the parabola in the third quadrant is approached, $e(x_0)$ emerges from this instability region, crosses the parabola and then enters the region again. (The gap contains the point $e(1/1)$, see below).

Location of Points Fixed under $(1/n)$

Periodic motion under T consists of 2 parts, one elliptic and the other hyperbolic, each for time $\tau/2$.

(a) Let the elliptic part intersect the x-axis at x_0, and the hyperbolic part intersect it at x_1. From Equations (7a) and (7b), their common points have $2\dot{x}^2 = (p_0/2)(x_0{}^4 - x_1{}^4)$. Consequently, *the orbit is symmetric about the x-axis*, and the fixed points lie on the x-axis, at x_0 and at x_1, respectively. If x_0 is positive and small enough so that $e(x_0)$ lies in the first quadrant ($x > 0$, $\dot{x} < 0$) then $he(x_0)$ will lie below the x-axis, since the hyperbolic motion is at lesser x than the elliptic motion and hence slower. Consequently $T(x_0) \neq x_0$ and x_0 cannot be a fixed point for T under these circumstances. If x_0 increases so that $e(x_0)$ is in the second quadrant, $he(x_0)$ still lies below the x-axis and x_0 is not a fixed point for T. When, however, x_0 is large enough so that $e(x_0)$ is in the third quadrant, the hyperbolic motion is at first toward the $(-x)$-axis and the time to go to this axis increases from zero, when $e(x_0)$ is on the $(-x)$-axis, to infinity when $e(x_0)$ is on the parabola $\dot{x} = (p_0/2)^{1/2}x^2$. For some value of x_0 in between, $he(x_0)$ will be on the $(-x)$-axis. There is then a fixed point (1/1) at this value of x_0. Actual calculation shows that $e(x_0)$ is not far from the parabola, so that we may take it at the parabola if we want a rough estimate for x_0. For such an intersection, $2x^4 = x_0{}^4$, or $cn\, u = \pm 2^{-1/4}$ and $u = 0.58778$, 3.12037, etc. Corresponding values of x_0 are 2.037, 10.814, 14.889, etc. The *fixed point* (1/1) therefore *lies a little below* $x_0 = 14.889$. The fixed points $(1/n)$ are located approximately at $x_0 = 2.037 + 12.852\, n$, $(n = 1, 2\,..)$.

(b) Let the elliptic part intersect the x-axis at x_0, and the hyperbolic part intersect the $\dot{x}$-axis as given by Equation (7c) with $x = 0$. From (7a) and (7c), the common points have $2x^4 = x_0{}^4 - \frac{1}{4}a^4$. *The orbit is symmetric about the $\dot{x}$-axis*, with fixed points at $\dot{x} = \pm(p_0/2)^{1/2}x_0{}^2$ and $\dot{x} = \pm(p_0/8)^{1/2}a^2$, for $n = 1$. A rough lower estimate, obtained by intersecting the ellipse with the parabola above, gives $x_0 = 2.037 + 6.426 = 8.463$. Thus, there is a fixed point (1/1) on the $\dot{x}$-axis lying somewhat above $\dot{x} = (p_0/2)^{1/2} x_0{}^2 = 9.74$. (Actual calculation shows it to be at about 10.26). The fixed points $(1/n)$ on the $\dot{x}$-axis are at values corresponding to $x_0 = 8.463 + (n-1)\ 12.852$.

More accurately, if a point P on the x-axis is fixed under T, then $he(P)$ must also lie on the x-axis, say at $x = x_1$ (positive or negative). Then, if we find the intersections of $e(x)$ with $h^{-1}(x)$ or $h^{-1}(-x)$, the fixed points are determined. The curve $e(x)$ spirals out from the origin, lying below $h^{-1}(x)$ in the first quadrant, and then intersecting $h^{-1}(-x)$ and $h^{-1}(x)$ alternately (third, fifth, etc. quadrants) and yielding fixed points $(1/n)$ with n a positive integer. At large distances h^{-1} is asymptotic to the parabolas P of Equation (15) or (15a). If the reflection of P in the x-axis is denoted by $\bar{P}$, then $e(\bar{P})$ will spiral around and intersect the x-axis near all the above fixed points $(1/n)$.

Conjecture

The above types of symmetry are found in general, and seem essential for periodic solutions of this differential equation. No rigorous proof has been devised, but our experience leads us to believe that *all the fixed points P of the equation* $\ddot{x} + px^3 = 0$, such that $T^n P = P$ (n = integer), *lie either on the x- or $\dot{x}$-axis, or are transforms* (TP, T^2P . . .) *of other* fixed points which do.

Successive Transformations $T^n(x)$

The curve $T(x)$ starts from the origin and spirals clockwise, intersecting the x-axis first at $-x_2$ and then at $+x_1$, with $|x_1| > |x_2|$. If x_a is the point which maps into $-x_2$, i. e. $T(x_a) = -x_2$, then $T^{-1}(x_a) = -x_2$, by symmetry, and $T^2(x_a) = T(-x_2) = x_a$. Consequently, x_a is a fixed point for T^2 as is $-x_2$. But $he(x)$ intersects $(he)^{-1}(-x)$ the first time on the $-\dot{x}$-axis, everything is symmetric, and so $x_a = x_2$. Two neighboring curves $T^n(x)$ and $T^{n+1}(x)$ will intersect when $T^{n+1}(x_b) = T^n(x_c)$, or $T(x_b) = x_c$. But then, by symmetry as above, x_b must be fixed under T^2 and $x_c = \pm x_b$. If we require that both x_b and x_c be positive, then they will be equal. Thus, $T^n(x)$ and $T^{n+1}(x)$, with $x > 0$, intersect for the first time at the fixed point (1/1).

Let us consider successive mappings Tx, T^2x, T^3x, where $0 \leq x \leq x_2$. Tx spirals clockwise to $-x_2$, T^2x spirals clockwise inside $T(x)$ to $+x_2$, T^3x inside T^2x to $-x_2$, etc. The operation T rotates the end of the curve by π each time. T^3x crosses the positive x-axis for some value $x_3 < x_2$, which is the fixed point $T^3x_3 = x_3$. Likewise, T^4x lies inside T^3x, and so the fixed

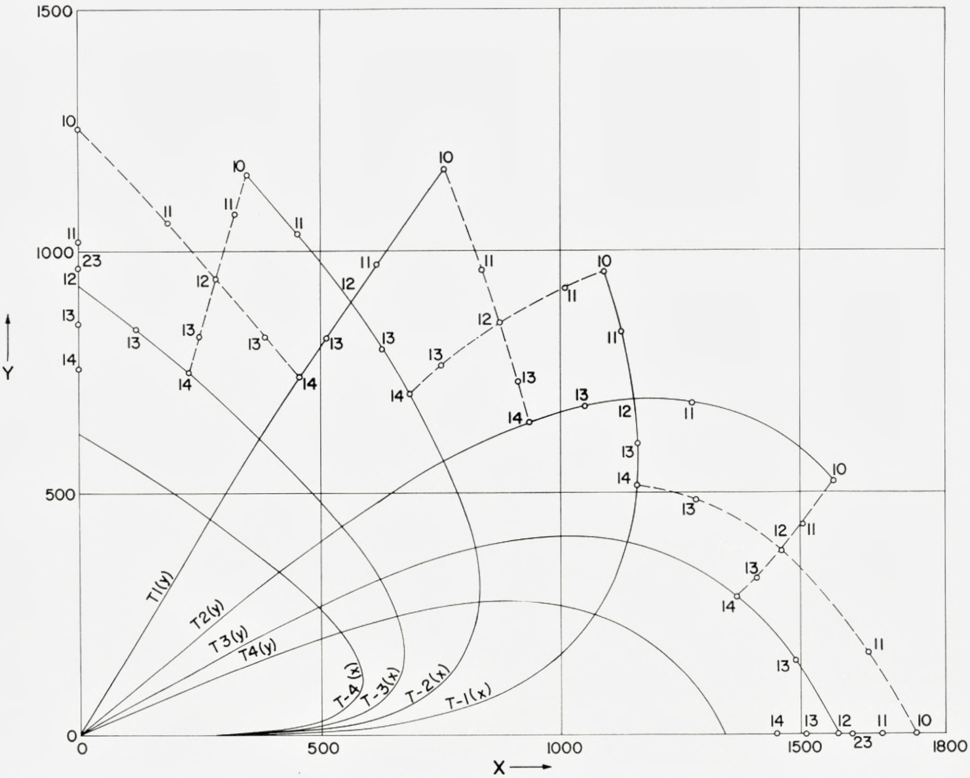

Figure 2. Fixed Points for $n = 10$ to $n = 14$ ($y = \dot{x}$ vs. x). Vertical Scale: $10^4\ \dot{x}$, Horizontal Scale: $10^3\ x$. Also mappings $T^{-n}(x)$ and $T^{n}(y)$ for $n = 1$ to $n = 4$.

point $T^4 x_4 = x_4$ will have $x_4 < x_3$. Thus we have an infinite number of isolated fixed points along the $+x$-axis, at least one for each value of n. (Similar reasoning applies to the $\dot{x}$-axis).

Now it can happen, as n becomes larger, that wrinkles develop in the curve $T^n x$ and that this will have more than one intersection with the $+x$-axis. Insertion of just one wrinkle will result in 3 intersections, i. e. $T^n(x_\alpha) = x_1$, $T^n(x_\beta) = x_2$, and $T^n(x_\gamma) = x_3$, where $x_\alpha < x_\beta < x_\gamma$. If these intersections are all distinct, then, since $T^n(x)$ cannot cross itself, either $x_1 < x_2 < x_3$ or $x_1 > x_2 > x_3$, In the first case, $x_1 = x_\alpha$, $x_2 = x_\beta$, $x_3 = x_\gamma$ and all the points are fixed under T^n. In the second case, $x_1 = x_\gamma$, $x_2 = x_\beta$, and $x_3 = x_\alpha$, so that only the middle point is fixed under T^n, and the outer points are fixed under T^{2n}.

Location of Points Fixed under (n/1)

A systematic search for points fixed under T^n can be made by looking at the mapping. First find where $T^n(x_a, \dot{x})$ intersects $x = x_a$ and plot the locus of such intersections. Then find where $T^n(x, \dot{x}_b)$ intersects $\dot{x} = \dot{x}_b$ and plot the corresponding locus. Intersections of the two loci give the fixed points. Such a search was made in the early stages of this work, and led to the observation that all the fixed points so obtained lay either on the x-axis or the $\dot{x}$-axis or were transforms of points which do.

Figure 2 shows, for $n = 10$, 11, 12, 13 and 14, where the fixed points ($x > 0$, $\dot{x} = y > 0$) lie, and in addition how the x and y axes map for $n = 1$, 2, 3, and 4. (Here T is defined as $e\,h^2\,e$). The point (23/2) is also shown.

A survey of some of the fixed points for low values of n was made by the Runge-Kutta method, for $p_0 = .037$ and period 6, and the results are tabulated below. If p is initially positive (first quarter-period), we deal with the basic transformation $T = e\,h^2\,e$, while if p is initially negative, the basic transformation is $T = h\,e^2\,h$.

TABLE I: Fixed points

x-axis, $T = e\,h^2\,e$

n	2	3	4	5	6	7	8
x	3.4143	3.1823	2.387	2.338	2.1086	2.058	1.9106
n	9	10	11	12			
x	1.8506	1.736	1.667	1.5777			

$\dot{x}$-axis, $T = e\,h^2\,e$

n	2	3	4	5	6	7	8
$\dot{x}$	9.6012	.4591	.3775	.3579	.3088	.1639	.1545
n	9	10	11	12			
$\dot{x}$	.1464	.1259	.1021	.09339			

x-axis, $T = h\,e^2\,h$

n	2	3	4	5	6	7	8
x	3.7968	2.1534	2.0564	1.6558	1.6276	1.5061	1.478
n	9	10	12				
x	1.4046	1.375	1.2926				

$\dot{x}$-*axis*, $T = h\, e^2\, h$

n	2	3	4	5	6	7	8
$\dot{x}$	1.355	1.082	.931	.2294	.2151	.2016	.1780
n	12						
$\dot{x}$	.09545						

Regions of Stability

Even though a study of the mapping near a fixed point immediately reveals whether this point is *elliptic or hyperbolic*, a further investigation is needed to ascertain whether or not an elliptic point is *stable*. ARNOL'D (1) and MOSER (2) claim that, if a mapping satisfies certain criteria near the fixed point, there exists an invariant curve around the fixed point such that no particle inside this curve can get outside; i. e. that the fixed point is stable. We shall now show how one can construct such a curve for the differential equation in question, if the curve exists.

Let us consider the points which are fixed under T^n. There will be a set of elliptic points ranged around the origin. Around each such point, if we are close enough so that the mapping is linear, any point P is mapped by T^n, T^{2n}, etc., into other points which lie on an ellipse. As the distance from the fixed point to P is increased, these ellipses swell out to meet corresponding ovals from neighboring fixed points. Their meeting place, roughly speaking, is marked by a hyperbolic fixed point. Thus, as we go round the origin, elliptic and hyperbolic points alternate, and the two pairs of characteristic lines from two neighboring hyperbolic points bend into invariant curves (I. C's) which enclose the elliptic point between them.

Let us now examine more closely the region between a hyperbolic point H_1 (for T^n) and its neighbor $T(H_1) = H_2$. Denote by α_1 and α_2 the I.C's for which the motion is away from H_1 and H_2 respectively, and by β_1 and β_2 the I.C's for which the motion is toward these points, as shown in Figure 3.

In practise, one can construct α_1 by taking points very close to H_1 and subjecting them repeatedly to the transformation T^n. The other invariant curves can be constructed in similar fashion.

Let α_1 and β_2 intersect at A_2, and α_2 and β_1 at B. The curve $H_1\ \alpha_1\ A_2\ \beta_2\ H_2\ \alpha_2\ B\ \beta_1\ H_1$ will be regarded as a bounding curve, and points inside which do not get outside under repeated transformations by T^n will comprise a *stable region* under T^n.

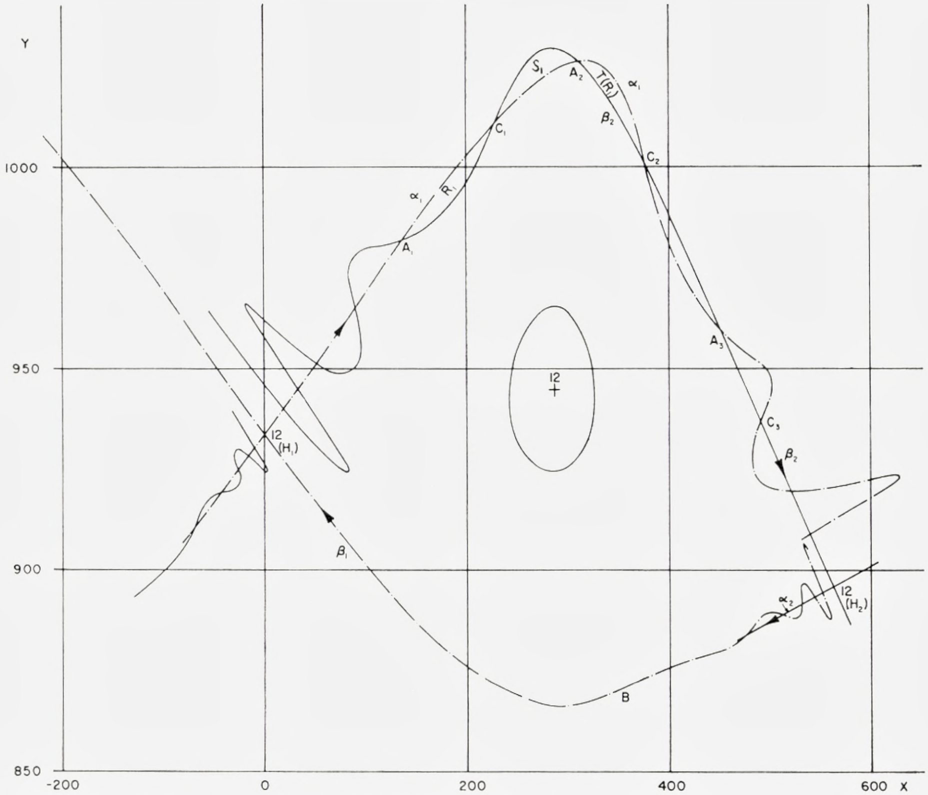

Figure 3. Invariant Curves from Hyperbolic Fixed Points H_1 and H_2 for $n = 12$ ($y = \dot{x}$ vs. x). An elliptic fixed point for $n = 12$, together with an approximate trajectory around it, is also shown. Vertical Scale: $10^4 \dot{x}$, Horizontal Scale $10^3 x$.

The nature of the invariant curve is such that oscillations of increasing amplitude develop. The point A_2 maps into $A_3 = T^n(A_2)$ on β_2. A point just inside A_2 must map into a point just inside A_3. Therefore, the curve α_1, which has crossed β_2 at A_2, must cross back between A_2 and A_3, say at C_2. Repeated application of T^n to A_3 gives a sequence of points A_4, A_5 etc., such that the distances $A_{n,n+1}$ are reduced each time by roughly a constant factor (the "stretch" factor mentioned earlier). Since area must be preserved, the amplitude of oscillation must increase each time to offset the compression against α_2. The end result is that *roughly one half of the points on* α_1 near H_1 "escapes", *and the other half approaches* α_2.

As β_2 proceeds backwards from A_2, it next crosses α_1 at $C_1 = T^{-n}(C_2)$ and then at $A_1 = T^{-n}(A_2)$. The region R_1 between β_2 and α_1, from A_1 to C_1, maps under T^n into a corresponding region between β_2 and α_1 from A_2 to C_2. Thus points inside the bounding curve have gotten outside, or escaped. *The complete region of escape is composed of* R_1 *and the corresponding area* R_2 *between* α_2 *and* β_1, *plus all of their maps under repeated application of* T^{-n}. At the same time that these particles are escaping, those from S_1 and S_2, and their maps under T^{-n}, are entering.

One may say that an I.C. has an inside and an outside, and that a point on the outside always stays there. However, points inside the I.C. α_1 can get outside β_2, as we have just seen. We have tacitly assumed that particles in $T(R_1)$ escape, when as a matter of fact they pile up on the I.C. from H_2 corresponding to α_1. But some of these stay outside and some go in, with the chances about even. There is continual fluctuation across the boundary, and in the long run, if it is just a question of T^n alone, no particles will escape from its sphere of influence, if they are initially inside the bounding curve.

If, however, a particle is in such an area as $T(R_1)$, which is outside the bounding curve β_2 for T^{11}, say, it may simultaneously be in a region of escape for T^{10}. Thus it may be handed on to outer regions until it hits our region of primary instability and escapes finally to infinity. The combined probability of leaking out altogether may not be very great. More to the point is the question as to how much fluctuation a particle will undergo in general, and whether or not it will cross the bounding curve for T^n.

The determination of how stable a particle is depends naturally on our definition of stability. If we wish to know what domains particles can reach from infinity, we need only apply $T^n e$ to the reflection of the curve of primary instability. The result of this, in our case, is that few particles penetrate in as far as the region of the T^{11} fixed points. As n increases, the probability gets smaller and smaller. If, however, we desire to know how many particles leave the domain primarily governed by T^n and come under the primary influence of T^{n-1}, we have merely to find the invariant curves of both transformations*, and then to see when the fluctuations, of the T^n I.C.'s are enough to intersect with the T^{n-1} I.C.'s.

This method does not lean at all on criteria of continuity as advanced by ARNOL'D and MOSER, but shows that the "invariant region" associated with a given transformation has as its boundary sections of the invariant curves issuing from the hyperbolic fixed points of this transformation.

* A similar remark holds for any T^m, $m \neq n$, but is of lesser significance when m is not close to n.

Acknowledgements

This work was conceived in 1956 at the Courant Institute for Mathematical Studies, while the author was on leave from the University of Illinois. It was supported at first by the University of Illinois Research Board, and later by the National Aeronautics and Space Administration (Grant No. NsG 280–63). The ideas on stability were crystallized during the summer of 1966, while the author was visiting the Copenhagen Observatory.

In addition to the previously mentioned discussion with Dr. J. Moser, there were valuable talks with Dr. D. J. Bourgin, who stressed the importance of ascertaining the mappings in the phase plane.

The author is very grateful for the above financial support and for the hospitality accorded him by Professor A. Reiz, Director of the Copenhagen Observatory, and by the USSR Academy of Sciences.

Department of Physics, University of Alabama,
University, Alabama.

References

(1) V. I. Arnol'd, Uspekhii mat. nauk, 18, 92–192 (1963).
(2) J. Moser, Nachrichten Akad. Wiss. Göttingen (Math-phys.) (1962), 1–20.
(3) G. Floquet, Ann. Ec. Norm. (2) 13, 47 (1883).
(4) J. L. Powell and R. S. Wright, MURA report RW/JLP – 4 (1955).
(5) J. Moser, Nachrichten Akad. Wiss. Göttingen (Math-phys.) (1955), 87–120.
(6) E. D. Courant and H. S. Snyder, Annals of Physics 3, 1–48 (1958).

Indleveret til Selskabet den 4. juli 1967.
Færdig fra trykkeriet den 16. februar 1968.