ON THE BEHAVIOR OF INVARIANT CURVES NEAR A HYPERBOLIC POINT OF A SURFACE TRANSFORMATION.*

By Shlomo Sternberg.

1. In [2], Poincaré considered surface transformations of the type

$$x_1 = sx + f(x, y), \qquad y_1 = ty + g(x, y), \tag{1}$$

where s and t are real with $s > 1 > t > 0$, and where f and g are real analytic functions of x and y which vanish at the origin together with their first derivatives. He proved, by the majorant method, that there exist two analytic invariant curves each tangent to one of the axes. In the case of dynamics, the relation $st = 1$ holds (as a necessary condition for the transformation to be area preserving). We shall not make use of this condition. Hadamard [1], treating the case where f and g are merely assumed to be of the class C^1 (and to vanish at the origin along with their first derivatives), showed the existence of invariant curves by the method of successive approximations. In paragraphs 2-6 we shall treat questions of existence and uniqueness of invariant curves by a geometric method under slightly less restrictive conditions and show that our results are, in a sense, the best possible. In paragraphs 7 and 8 we shall deal with questions of smoothness of invariant curves and show that a C^n assumption on the non-linear terms implies that the invariant curves are of class C^n $(n \geqq 1)$. The problems treated in this paper were suggested to the author by Professor Wintner.

2. Let T be a transformation of type (1) which is topological in some neighborhood of the origin. Let $s > 1 > t > 0$, and let the functions f, g be continuous and $o(r)$ as $r = (x^2 + y^2)^{\frac{1}{2}} \to 0$. It is then clear that for all points sufficiently close to the origin, the y coordinate is decreased and the x coordinate is increased upon application of T. Thus if an arc passing through the origin is to be invariant, it must be tangent to one of the axes or else experience oscillations of increasing amplitude and frequency near the origin. In the latter case we shall, naturally, not speak of an "invariant curve." Hence, either the arc is tangent to the x-axis at the origin, in which case all points of the arc move in toward the origin upon application of T; or

* Received February 7, 1955.

the arc is tangent to the y-axis, in which case all the points move away from the origin.

3. In this paragraph and the following, S and T will denote transformations of type (1) where $t > 1$, $t > s$ and $s > 1$, $s > t$ respectively.

THEOREM 1. *Let S be a transformation of type* (1) *where* $0 < s < 1$, $s < t$ *and where* f, g *are continuous and are* $o(r)$ *as* $r = (x^2 + y^2)^{\frac{1}{2}} \to 0$. *Then there exist a neighborhood N of the origin and a closed set E in N such that* 1) *E is invariant under S,* 2) *for every abscissa x in N there exists a point (x, y) of E,* 3) *E is tangent to the x-axis of the origin* (i.e., given any $\epsilon > 0$, for all sufficiently small $|x|$, all points (x, y) of E lie in the angular region $|y| \leqq \epsilon |x|$).

Proof. We can restrict our attention to the right half plane $x \geqq 0$ since the argument is the same for the left half plane. Consider a point (x, y) on the line $y = \alpha x$, $\alpha > 0$. Its image will be the point (x_1, y_1), where $x_1 = sx + f(x, y)$, $y_1 = ty + g(x, y)$. Thus

$$y_1/x_1 = \{\alpha t/s + g(x, y)/sx\}/\{1 + f(x, y)/sx\}.$$

Since $y = \alpha x$ and both f and g are $o(r)$, we find that $x_1 < x$ and $y_1 > \alpha x_1$ for all sufficiently small positive x, say, for $0 < x \leqq x_0$.

In other words, a point on the line $y = \alpha x$ is moved to the left and above the line by S for all $x < x_0$. Similarly, a point on the line $y = -\alpha x$ is moved to the left and below the line. For all $x < x_0$, let $E^0(x)$ denote the vertical line segment joining the points $(x, -\alpha x)$ and $(x, \alpha x)$. Now S will map $E^0(x)$ into a set containing a connected arc joining a point above the wedge $|y| \leqq \alpha x$ to a point below the wedge. Let $E^1(x)$ denote the subset of $E^0(x)$ consisting of those points whose images under S lie in the wedge $|y| \leqq \alpha x$. If $E^0(x), E^1(x), \cdots, E^{n-1}(x)$ have been defined, let $E^n(x)$ denote the subset of $E^{n-1}(x)$ consisting of those points whose images under S^n lie in the wedge. Since, by induction, the image of $E^{n-1}(x)$ under S^n contains a connected arc joining a point above the wedge to a point below it, $E^n(x)$ is not empty. $E^n(x)$ is clearly closed for all n. Hence $E(x) = \bigcap_n E^n(x)$ is a non-empty closed subset of $E^0(x)$. It is clear that given any ϵ we can find a positive number ξ such that for all $0 < x \leqq \xi$ and all $k \geqq \epsilon$, S takes a point on the line $y = kx$ into a point above the line $y = (k + \tau)x$ where τ is independent of x and k. Hence for a point (x, y) with $x \leqq \xi$ to lie in E we must have $|y| < \epsilon x$. Thus $E(x)$ is tangent to the x-axis at the origin.

Let $E = \bigcup_x E(x)$. Then E is clearly invariant under S. In fact, any subset of the wedge invariant under S is eventually contained in E. Furthermore, a point $P = (x, y)$ of the closed wedge $|y| \leqq \alpha x$, $0 \leqq x \leqq x_0$ is not in E only if S^n, for some positive n, carries P outside the closed wedge. But then a small neighborhood about P is carried outside the wedge by S^n. Thus the complement of E is open, so that E is closed. This proves 1), 2) and 3).

THEOREM 2. *Let T be a transformation of type* (1) *where $s > 1$, $s > t$ and f and g are continuous and $o(r)$ as $r \to 0$. Then there exist a neighborhood N of the origin and a closed set E satisfying* 1), 2), *and* 3) *of Theorem* 1 *where, by invariance, we mean $T(E) \cap N = E$.*

Remark. If T is topological, then Theorem 2 can be obtained by applying Theorem 1 to $S = T^{-1}$.

In order to prove Theorem 2, let us consider an angular wedge W: $|y| \leqq kx$. Then from arguments similar to the above we have

$$(T(W) \cap N) \subset (W \cap N).$$

Let

$$F_0 = W \cap N, \qquad F_1 = T(W \cap N) \cap N$$

and in general $F_n = T(F_{n-1}) \cap N$. It is clear, by induction, that $F_{n+1} \subset F_n$ and thus $E = \bigcap_n F_n$ is a non-empty closed set. It is clear from arguments similar to those of the preceding theorem that 1), 2) and 3) are satisfied.

4. We shall now show that, under the hypotheses of Theorem 1, $E(x)$ can contain more than one point. S will be a transformation of type (1) in which $f \equiv 0$. We shall choose g such that $g(x, 0) = 0$ and so that the x-axis is an invariant curve.

Let $y = \psi(x)$ be a smooth curve tangent to the x-axis at the origin and such that $\psi(x) > 0$ for $x > 0$. In order that $\psi(x)$ be an invariant curve, it is necessary and sufficient that the functional equation

$$t\psi(x) + g(x, \psi(x)) = \psi(sx) \tag{2}$$

be satisfied. Rewriting this we obtain the condition on g

$$g(x, \psi(x)) = \psi(sx) - t\psi(x). \tag{2'}$$

We now define g as follows: Let $g(x, y)$ be 0, $\{\psi(sx) - t\psi(x)\}y/\psi(x)$ or $\psi(sx) - t\psi(x)$ according as $y \leqq 0$, $0 \leqq y \leqq \psi(x)$ or $y \geqq \psi(x)$. We must now verify that the mapping S, with $f \equiv 0$ and g defined as above, satisfies the conditions of Theorem 1. (It is clear that ψ can be so chosen that the mapping is even topological.) Now $|g(x,y)| \leqq \psi(sx) + t\psi(x)$. Thus g is $o(|x|)$ and so *a fortiori* $o(r)$.

We have constructed g so that (2′) is satisfied; hence, both $y = \psi(x)$ and $y = 0$ are invariant curves of S.

5. We now consider the question of uniqueness.

THEOREM 3. *If the hypotheses of Theorem 1 are satisfied, and if both f and g satisfy uniform Lipschitz conditions with respect to x and y, where in a neighborhood of the origin, the uniform Lipschitz constants are less than $\delta < \frac{1}{4}(t - s)$, then, in a suitably small neighborhood N, E is a curve $y = \psi(x)$. In other words, there exists a unique invariant arc in any wedge $|y| \leqq k|x|$ for sufficiently small $|x|$.*

In order to prove this theorem it will suffice to show that given any wedge $|y| \leqq kx$ there exists a positive number ξ such that there cannot be two points (x,y) and (x,y'), where $0 < x \leqq \xi$ and $y > y'$, whose images under S^n are in $|y| \leqq kx$ for all non-negative n.

Since f is $o(r)$, any point in $|y| \leqq kx$ with sufficiently small abscissa $x\,(>0)$ is mapped into a point whose abscissa is smaller than $(s+\epsilon)^n x$. Thus if (x_n, y_n) and (x'_n, y'_n) denote the images under S^n of (x,y), (x,y'), we must have $x_n, x'_n < (s+\epsilon)^n x$ and consequently, $|y_n - y'_n| < 2k(s+\epsilon)^n x$ in order that these images all lie in $|y| \leqq kx$. We shall show that this is impossible for all suitably small x.

From (1), $x_n - x'_n = s(x_{n-1} - x'_{n-1}) + f(x_{n-1}, y_{n-1}) - f(x'_{n-1}, y'_{n-1})$. Thus, for all sufficiently small x,

$$|x_n - x'_n| \leqq s\,|x_{n-1} - x'_{n-1}| + \delta\,|x_{n-1} - x'_{n-1}| + \delta\,|y_{n-1} - y'_{n-1}|.$$

Similarly, $y_n - y'_n = t(y_{n-1} - y'_{n-1}) + g(x_n, y_n) - g(x'_n, y'_n)$; so that

$$|y_n - y'_n| > t\,|y_{n-1} - y_{n-1}| - \delta\,|x_{n-1} - x'_{n-1}| - \delta\,|y_{n-1} - y'_{n-1}|.$$

Now assume that

$$(3) \qquad |x_{n-1} - x'_{n-1}| < |y_{n-1} - y'_{n-1}|;$$

then we have

$$|x_n - x'_n| < (s + 2\delta)|y_{n-1} - y'_{n-1}| \quad \text{and} \quad |y_n - y'_n| > (t - 2\delta)|y_{n-1} - y'_{n-1}|.$$

Since $\delta < \frac{1}{4}(t-s)$ we obtain $|x_n - x'_n| < |y_n - y'_n|$. As $|x_0 - x'_0| = 0 < |y_0 - y'_0|$, the inequality (3) is true for all n, by induction. We thus obtain

$$|y_n - y'_n| > (t - 2\delta)|y_{n-1} - y'_{n-1}| \quad \text{or} \quad |y_n - y'_n| > (t - 2\delta)^n\,|y - y'|.$$

We must therefore have $(t-2\delta)^n\,|\,y-y'\,|<2k(s+\epsilon)^n x$, which is clearly impossible for all n if ϵ is chosen to be less than $\frac{1}{2}(t-s)$. Thus $y=y'$ and the theorem is established.

6. Let us now consider a transformation T satisfying the hypotheses of Theorem 2. The set E, in virtue of its construction, contains any invariant set lying in the intersection of any angular wedge $|\,y\,|\leqq kx$ and some sufficiently small neighborhood of the origin. Now let G_0 be any set lying entirely in such a region. It follows, from reasoning similar to that of paragraph 3, that the image of this set under T, G_1, still lies in the angular region $|\,y\,|\leqq kx$. Let $G_0, G_1, G_2, \cdots, G_n$ be the sequence of sets so obtained. Let C be the set of limit points of this sequence. C is obviously an invariant set and hence a subset of E. In particular, if E reduces to a curve, then all the successive approximations converge to the invariant curve. Let us now assume that both f and g satisfy uniform Lipschitz conditions (with respect to x and y) where the Lipschitz constant can be made arbitrarily small by restricting ourselves to a small neighborhood of the origin. Let $y=\phi_0(x)$ be a curve satisfying a uniform Lipschitz condition with Lipschitz constant K. Thus $|\,(y-y')/(x-x')\,|<K$. Using the same notation as before, we obtain

$$(y_1-y'_1)/(x_1-x'_1)$$
$$=\{t(y-y')+g(x,y)-g(x',y')\}/\{s(x-x')+f(x,y)-f(x',y')\},$$

so that $|(y_1-y'_1)/(x_1-x'_1)|<\{tK+(1+K)\delta\}/\{s-(1+K)\delta\}$, where δ can be made arbitrarily small by choosing x and x' sufficiently small. Thus in a sufficiently small region about the origin, all the image curves satisfy a Lipschitz condition with Lipschitz constants all smaller than K. Hence the sequence is equi-continuous and the convergence is uniform. This is essentially the case considered by Hadamard [1]. We have thus proved

THEOREM 4. *Let T be a transformation of type* (1) *where $s>1$, $s>t$, f and g are continuous and $o(r)$. Furthermore let the set E of Theorem 2 be a curve. Then the sequence of successive images of a set situated in an angular region $|\,y\,|<kx$ converge for sufficiently small x. If, in addition, f and g satisfy uniform Lipschitz conditions with Lipschitz constants which are $o(1)$ as $r\to 0$, then the iterates of any curve $y=\phi(x)$ which satisfies a uniform Lipschitz condition converge uniformly for sufficiently small x to the invariant curve, which then must satisfy a uniform Lipschitz condition.*

7. We shall now consider differentiability of the invariant curve.

THEOREM 5. *Let S be a transformation of type (1), where $0 > s > t$, $s > 1$ and where f and g are of class C^1 and vanish together with their first derivatives at the origin. Then the unique invariant curve tangent to the x-axis at the origin is of class C^1 for all sufficiently small x.*

In order to make the method clear, we shall first go through the proof for the particular case when $f(x,y) \equiv 0$. Let $y = \psi(x)$ be the invariant curve. It must then satisfy the functional equation (2). Writing this equation for two distinct points x and x' and subtracting, we obtain

$$t\psi(x) - t\psi(x') + g(x,\psi(x)) - g(x',\psi(x')) = \psi(sx) - \psi(sx').$$

Applying the mean value theorem, we obtain

$$t\{\psi(x) - \psi(x')\}/(x - x') + g_x(\xi^*,\psi(x)) + g_y(x',\psi(\xi))\{\psi(x) - \psi(x')\}/(x - x')$$
$$= \{\psi(sx) - \psi(sx')\}/(x - x') \text{ or}$$

$$\{\psi(x) - \psi(x')\}/(x - x')$$
$$= [t + g_y(x',\psi(\xi))]^{-1}[-g_x(\xi^*,\psi(x)) + s\{\psi(sx) - \psi(sx')\}/(sx - sx')],$$

where both ξ and ξ^* lie between x and x'.

We have thus expressed $\{\psi(x) - \psi(x')\}/(x - x')$ in terms of

$$\{\psi(sx) - \psi(sx')\}/(sx - sx').$$

Repeating the process n times, we obtain

$$\{\psi(x) - \psi(x')\}/(x - x')$$
$$= -\sum_{k=1}^{n} s^{k-1} g_x(\xi^*_{k-1},\psi(s^{k-1}x))\{\prod_{j=1}^{k}(t + g_y(s^{j-1}x,\psi(\xi_{j-1}))\}^{-1}$$
$$+ s^n\{\prod_{j=1}^{n}(t + g_y(s^{j-1}x,\psi(\xi_{j-1}))\}^{-1}\{\psi(s^n x) - \psi(s^n x')\}/(s^n x - s^n x'),$$

where both ξ_j and ξ^*_j lie between $s^j x$ and $s^j x'$.

Now $\psi(x)$ satisfies a uniform Lipschitz condition (by Theorem 4) and $|g_y| < t - s - \delta$ for all sufficiently small x and x'. Thus the second term in the above equation tends to zero and we obtain

$$\{\psi(x) - \psi(x')\}/(x - x') = \sum_{k=1}^{\infty} s^k g_x(\xi^*_{k-1},\psi(s^{k-1}x)) \prod_{j=1}^{k}(t + g_y(s^{j-1}x,\psi(\xi_{j-1})))^{-1}.$$

Since g_x and g_y are $o(1)$ as $r \to 0$, this series is majorized by a series of the form $\sum_k s^{k-1}(t-\epsilon)^{-k}$ and so is uniformly convergent. We can, therefore, let $x \to x'$ and obtain

$$\psi'(x) = \sum_{k=1}^{\infty} s^{k-1} g_x(s^{n-1}x, \psi(s^{k-1}x)) \prod_{j=1}^{k} (t + g_y(s^{j-1}x, \psi(s^{j-1}x)))^{-1}.$$

In the case where f is not identically zero the iteration is more difficult. The functional equation satisfied by the invariant curve $\psi(x)$ now takes the form

$$t\psi(x) + g(x, \psi(x)) = \psi(sx + f(x, \psi(x))).$$

(This is the functional equation used by Poincaré [2]). Proceeding as before,

$$t\{\psi(x) - \psi(x')\} + g(x, \psi(x)) - g(x', \psi(x')) = \psi(x_1) - \psi(x'_1),$$

where $(x_1, \psi(x_1))$ and $(x'_1, \psi(x'_1))$ are the images, respectively, of $(x, \psi(x))$ and $(x', \psi(x'))$. Thus

$$\begin{aligned} t\{\psi(x) - \psi(x')\}/(x - x') + g_x(\xi^*, \psi(x)) + g_y(x', \psi(\xi))\{\psi(x) - \psi(x')\}/(x - x') \\ = [\{\psi(x_1) - \psi(x'_1)\}/(x_1 - x'_1)][s + f_x(\eta^*, \psi(x)) \\ + f_y(x', \psi(\eta))\{\psi(x) - \psi(x')\}/(x - x')], \end{aligned}$$

where ξ, ξ^*, η and η^* are all between x and x'. Hence, the iteration takes the form $[\ ]_0 = \{(s + \sigma_0)[\ ]_1 + \delta_0\}\{\epsilon_0[\ ]_1 + (t + \tau_0)\}^{-1}$, where

$$[\ ]_0 = \{\psi(x) - \psi(x')\}/(x - x'), \qquad [\ ]_1 = \{\psi(x_1) - \psi(x'_1)\}/(x_1 - x'_1),$$

$$\sigma_0 = f_x(\eta^*, \psi(x)), \quad \delta_0 = -g_x(\xi^*, \psi(x)), \quad \epsilon_0 = f_y(x', \psi(\eta)), \quad \tau_0 = g_y(x', \psi(\xi)).$$

Introducing corresponding notations for the k-th iterate of this relation, we obtain

$$[\ ]_0 = \prod_{j=0}^{k} \begin{pmatrix} s + \sigma_j & \delta_j \\ \epsilon_j & t + \tau_j \end{pmatrix} [\ ]_{k+1}, \text{ where } \begin{pmatrix} a & b \\ c & d \end{pmatrix}(z) = (az + b)/(cz + d).$$

and the product denotes matrix multiplication.

We now wish to establish the convergence of this iteration. We first show that for all sufficiently small x and x', the product

$$\prod_{j=0}^{k} \begin{pmatrix} s + \sigma_j & \delta_j \\ \epsilon_j & t + \tau_j \end{pmatrix}$$

can be put in the form

$$\begin{pmatrix} A_k & B_k \\ C_k & 1 + D_k \end{pmatrix}, \text{ where } A_k = A \prod_1^k \alpha_j,\ C_k = C \prod_1^k \gamma_j,$$

with $|\alpha_j| < r < 1$, $|\gamma_j| < r < 1$, and $B_k = B_{k-1} + \beta_k$, $D_k = D_{k-1} + \partial_k$, with

$|\beta_k| < p^k$, $|\partial_k| < p^k$, and $|p| < 1$. This is certainly true for the case $k = 0$. We shall prove that the matrix has this form for all n, by induction. Thus assume the proposition true for the case $k = n$. Then the $(n+1)$-st matrix is

$$\begin{pmatrix} A_n & B_n \\ C_n & 1+D_n \end{pmatrix} \begin{pmatrix} s+\sigma_n & \delta_n \\ \epsilon_n & t+\tau_n \end{pmatrix}.$$

Now since this matrix product is considered as acting as a linear fractional transformation we may divide all four elements by the same amount without changing the effect of the matrix. Carrying out the matrix multiplication and dividing by $t + \tau_n$ we obtain the matrix

$$\begin{pmatrix} A_n(s+\sigma_n)\rho_n + B_n\epsilon_n\rho_n & A_n\delta_n\rho_n + B_n \\ C_n(s+\sigma_n)\rho_n + (1+D_n)\epsilon_n\rho_n & 1 + D_n + C_n\delta_n\rho_n \end{pmatrix},$$

where $\rho_n = 1/(t+\tau_n)$. Now by virtue of the inductive hypothesis, $|B_n| < B$ and $|D_n| < D$, where B and D do not depend on n. Since σ_n, δ_n, ϵ_n, τ_n are all uniformly $o(1)$ (as x and x' go to zero), we may choose x and x' so small that

$$|(s + \sigma_n + \epsilon_n B/A)/(t+\tau_n)| < r < 1,$$

$$|(s + \sigma_n + \epsilon_n(1+D_n)/C)/(t+\tau_n)| < r < 1,$$

$$|A\delta_n/(t+\tau_n)| < p < 1, \qquad |C\delta_n/(t+\tau_n)| < p < 1.$$

Thus the $(n+1)$-st matrix has the desired form. It should be remarked at this point that the constants C and D are still at our disposal.

As before, we know that the $[\ ]_n$ are uniformly bounded. We now choose C and D so small that the denominator of the n-th iterate never vanishes and the uniform convergence of the iteration process is established.

COROLLARY. *If, in addition to the hypotheses of Theorem 5, f and g are assumed to be of class C^n, then ψ is of class C^n in some neighborhood of the origin.*

Proof. The function $\psi'(x)$ can be represented as the quotient of two series both of which are majorized by $\sum Kr^n$ with $r < 1$. In taking successive difference quotients, one obtains series majorized by $K\sum M^k(n!/(n-k)!)r^n$, which is also uniformly convergent.

8. Let T be a transformation of type (1), where f and g are of class C^n $(n \geqq 1)$ and vanish at the origin together with their first derivatives. Furthermore, let $s > 1 > t > 0$. Then both T and T^{-1} satisfy the conditions of the corollary to Theorem 5 and hence, there exist exactly two invariant curves of the transformation T, each tangent to one of the axes at the origin. We can perform a rotation of the coordinate system so that now neither invariant curve touches an axis. In these new (u, v) coordinates the two invariant curves may be represented as $v = \phi_1(u)$ and $v = \phi_2(u)$. We now write $X = v - \phi_1(u)$, $Y = v - \phi_2(u)$. Since the curves ϕ_1 and ϕ_2 are perpendicular at the origin, the Jacobian of this transformation does not vanish. The transformation T is therefore defined in these new variables. We thus have

THEOREM 6. *Let T be a transformation of type* (1) *with $s > 1 > t > 0$. Furthermore, let f and g be of class C^n $(n \geqq 1)$ and vanish at the origin together with their first derivatives. Then there exists a non-singular change of coordinates $X = X(x, y)$, $Y = Y(x, y)$ of class C^n so that in the new coordinates the axes are the invariant curves and T has the form*

$$X_1 = s(X + F(X, Y)), \qquad Y_1 = t(Y + G(X, Y)),$$

where F and G are both of class C^n, and where

$$F(0, y) = G(x, 0) = F_x(0, 0) = F_y(0, 0) = G_x(0, 0) = G_y(0, 0) = 0.$$

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] J. Hadamard, *Selecta*, 1901, pp. 163-166.

[2] H. Poincaré, Œuvres, vol. 1 (1886), pp. 202 204