A GRAPH THEORETIC PROOF OF SHARKOVSKY'S THEOREM ON THE PERIODIC POINTS OF CONTINUOUS FUNCTIONS

CHUNG-WU HO AND CHARLES MORRIS

Let f be a continuous real valued function defined on the real line. If f has a periodic point of period k, does f have to have a periodic point of some other period m? A Russian mathematician, A. N. Sharkovsky obtained a complete answer to this question. Sharkovsky's result is elegant, however, his proof is difficult. Recently, P. D. Straffin attempted to give a simple proof of the sufficient part of Sharkovsky's theorem by means of directed graphs. However, his proof contains a gap. In this paper, the authors fill in the gap in Straffin's work. They also give a proof of the necessary part of the theorem, which is also based on directed graphs, and thus, obtain a complete simple proof of Sharkovsky's theorem.

I. **Introduction.** Let $f: R \to R$ be a continuous function. If f has a periodic point of a period k, must f also have points of other periods $m \neq k$? In 1975, T.-Y. Li and J. Yorke published a surprising result on this question: If $f: R \to R$ has a point of period 3, then it has points of all other periods [3]. However, unknown to Li and Yorke, a Russian mathematician, A. N. Sharkovsky had already published a more surprising result in 1964; namely, a complete answer to the question ([4], see also [5]). Arrange the poitive integers into the following sequence:

$$3, 5, 7, 9, \cdots, 3 \cdot 2, 5 \cdot 2, 7 \cdot 2, \cdots, 3 \cdot 2^2, 5 \cdot 2^2, \cdots, \cdots, 2^3, 2^2, 2, 1 \, .$$

Sharkovsky's theorem says that a function $f: R \to R$ having a point of period m, must also have points of period n precisely when m precedes n in the above sequence.

Sharkovsky's proof, however, is long and complicated. In the words of P. D. Straffin, "He constructs so many sequences of points that eight complex figures and most of the letters of the Greek alphabet are necessary to keep track of them." ([6, p. 104]). Straffin attempted to give a simpler proof of the sufficient part of Sharkovsky's theorem, i.e., period m implies period n if m precedes n in the above sequence. Straffin's method is interesting. Corresponding to a periodic point of f, he constructed a digraph (i.e., a directed graph). He showed that the existence of certain cycles of length k in the digraph implies the existence of periodic points of a period k for f [6]. However, Straffin did not quite succeed in proving the sufficient part

of Sharkovsky's theorem. His proof contains a gap. In this paper, we fill in the gap described by Straffin. Again using digraph, we shall also prove that the Sharkovsky's sequence is strict, and thus, obtain a complete simple proof of Sharkovsky's theorem. We are informed recently that F. Fuglister has also found an independent proof of the theorem, and that the digraphs of our Figure 4 have also appeared in a summary of Straffin's work in [1, Chapt. 6]. Besides, we would like to point out that Sharkovsky's theorem can also be proved for a certain class of smooth functions by means of symbolic dynamics (see [2]). In §II below, we first summarize Straffin's main results. Our proof of Sharkovsky's theorem will be presented in §III. For completeness, part of the standard argument used by previous authors will also be repeated here. Finally, a few observations and a question will be listed in §IV.

II. **A summary of Straffin's results.** Let $f: R \to R$ be a continuous function and x be a periodic point of f of period k. Let a be the smallest number among x and its iterates under f. Mark these points on the real number axis. They divide the real line into two infinite rays and $k - 1$ finite intervals. These finite intervals are to be labeled as $I_1, I_2, \cdots, I_{k-1}$ from left to right. Let s_i and t_i be the end points of the interval I_i. We say that there is a *direction* from the interval I_i to the interval I_j if I_j lies in between the two points $f(s_i)$ and $f(t_i)$. A directed graph can then be formed by using vertices corresponding to these intervals. The vertices are again labeled $I_1, I_2, \cdots, I_{k-1}$. A directed arc is drawn from the vertex I_i to the vertex I_j if there is a direction from the interal I_i to the interwal I_j[1]. This directed graph is called a *k-periodic digraph* of f. As an example, a function with a periodic point $a = 0$ of period 4, and the corresponding 4-periodic digraph are given in Figures 1 and 2 respectively.

A cycle in the periodic digraph is said to be a *nonrepetitive cycle* if it does not consist *entirely* of a cycle of smaller length traced several times. For instance, the cycle $I_1I_3I_1I_3I_1$ in Figure 2 is a repetitive cycle of length 4, while $I_1I_3I_1$ is a nonrepetitive cycle of length 2. The usefulness of digraphs is demonstrated by the following theorem of Straffin [6].

THEOREM 2.1. *If a k-periodic digraph of f has a nonrepetitive cycle of length m, then f has a periodic point of period m.*

---

[1] Our definition is slightly different from that of Straffin. He requires the graph to have a directed arc I_iI_j if $f(I_i) \supset I_j$. This difference can be seen for the function in our Figure 1. His definition would require an arc from I_2 to I_1. However, it appears that Straffin was actually using a definition equivalent to ours; for otherwise his enumeration of all 5-periodic digraphs [6, p 13] would not be complete.
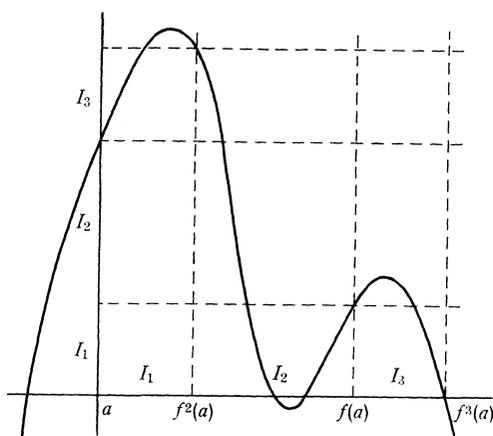
FIGURE 1
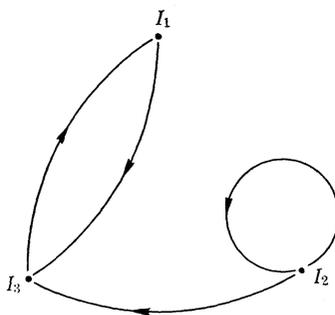
FIGURE 2

It is easy to show that the converse of this theorem is not true in general. However, a partial converse will be proved in §III (see Proposition 3.5). Using this theorem, Straffin gave a very simple proof of Li and Yorke's result that period $3 \Rightarrow$ period m for all m. He then established the following

THEOREM 2.2. *If a continuous function* $f: R \to R$ *has a point of odd period* $k > 1$, *then it must have periodic points of all periods greater than or equal to* $k - 1$.

In proving this theorem, Straffin used an idea which we shall use later. For ease of reference, we state this idea in terms of the following

LEMMA 2.3. *A* k-*periodic digraph of a continuous function always contains a cycle of length* k *in which some vertex is repeated exactly twice.*

The cycle of length k constructed by Straffin in general may be a repetitive cycle (e.g., Straffin's method, if applied to the function in our Figure 2, yields a repetitive 4-cycle). However, since this k-cycle contains a vertex I exactly twice, it can always be decomposed into two smaller cycles, each of which contains I exactly once, and therefore, both of the smaller cycles are nonrepetitive.

III. **Proof of Sharkovsky theorem.** The theorem will be proved by a sequence of propositions.

PROPOSITION 3.1. *If a continuous function* $f: R \to R$ *has a point of odd period* $2n + 1 (n \geq 1)$, *then it must have periodic points of all even periods.*

Proof. When $n = 1$, the assertion follows from Li and Yorke's result. We shall consider the case when $n > 1$. First construct a $2n + 1$ periodic digraph of f. By Lemma 2.3 and following comments, the digraph has a $(2n + 1)$ cycle, which decomposes into two smaller nonrepetitive cycles, one of which has an odd length. If this odd length cycle has a length greater than 1, the proposition is immediately proved by induction. Hence, we may assume that the $(2n + 1)$ cycle decomposes into a 1 cycle and a $2n$ cycle. Let J_1 be the vertex connecting the 1 cycle and the $2n$ cycle, and label the $2n$ cycle, starting with J_1, by $J_1 J_2 \cdots J_{2n} J_1$.

We may assume that the $2n$ cycle does not contain any proper subcycle of length greater than 1 which contains J_1 as a vertex, for otherwise, we may choose a nonrepetitive such subcycle, and then either add or not add the 1 cycle to this subcycle to get a nonrepetitive cycle of an odd length which is strictly between 1 and $2n + 1$, and thus, complete the proof by induction. Consequently, we may assume that in the $2n + 1$ periodic digraph, each vertex $J_i (1 < i < 2n)$ cannot be directed to any vertices J_k for $k > i + 1$, or for $k = 1$, for otherwise, we may bypass a few vertices and obtain a proper subcycle of length greater than 1 which contains J_1 as a vertex. Similarly, because of the 1 cycle at J_1, the vertex J_1 is directed only to J_1 and J_2, but not to any other vertices.

Let $J_1 = I_m$, the mth finite interval in the partition of the real line by the images of the periodic point (see §II). Since J_1 is directed only to J_1 and J_2, J_2 must be an interval just next to I_m, and f must carry one of the end points of J_1 into an end point of J_1, and the other end point of J_1 into an end point of J_2. Let b be the end point of J_1 such that $f(b)$ is again an end point of J_1. Since the point b is of the period $2n + 1$, $f(b) \neq b$. Therefore, there are only two possibilities:

(1) $J_1 = [b, f(b)]$, and in this case, $J_2 = I_{m-1} = [f^2(b), b]$.
(2) $J_1 = [f(b), b]$, and $J_2 = I_{m+1} = [b, f^2(b)]$.
By symmetry, we shall treat only the first case.

$$\begin{array}{ccc} & J_2 & J_1 \\ \hline f^2(b) & b & f(b) \end{array}$$
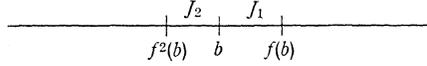
Note that $f^3(b) > f(b)$, for otherwise, we would have $f^3(b) < f^2(b)$, and therefore, J_2 would have to be directed to J_1. Also since J_2 is not directed to any J_k for $k > 3$, we conclude that J_3 lies to the right of J_1, and $J_3 = [f(b), f^3(b)]$. Similarly, $f^4(b) < f^2(b)$ and $J_4 = [f^4(b), f_2(b)]$. Continuing this argument, we conclude that the intervals J_i, $1 \leqq i \leqq 2n$, must be located in the real line as follows:

$$\begin{array}{ccccccccc} & J_{2n} & \cdots & J_4 & J_2 & J_1 & J_3 & \cdots & J_{2n-1} \\ \hline f^{2n}(b) & f^{2n-2}(b) & \cdots & f^4(b) & f^2(b) & b & f(b) & f^3(b) & \cdots & f^{2n-3}(b) \quad f^{2n-1}(b) \end{array}$$
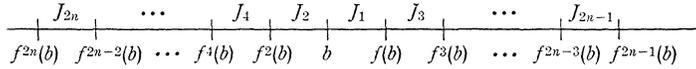
where $f^{2n+1}(b) = b$ and J_{2n} is directed to $J_1, J_3, \cdots, J_{2n-1}$. In particular, the $2n + 1$ periodic digraph must contain the subgraph
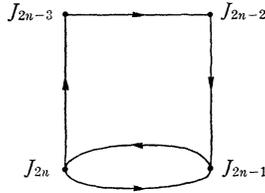
from which the existence of all even periods is clear.

PROPOSITION 3.2.   *If a continuous function $f: R \to R$ has a point of any period k, it must have a point of period 1.   Also, if f has a point of any period $k > 1$, then it has a point of period 2.*

Proof.   We shall only prove the second assertion. The first one can be proved in a similar manner (for an alternate simple proof of the first assertion see [6, p. 99]). The statement is trivially true for $k = 2$. Now suppose $k \geqq 3$. By Lemma 2.3, the corresponding k-periodic digraph of f contains a k cycle, which decomposes into two nonrepetitive subcycles, at least one of which has a length greater than 1. Thus, the statement follows by induction.

PROPOSITION 3.3.   *Let $f: R \to R$ be a continuous function.   If h precedes k in the Sharkovsky's sequence, and if f has a point of period h, then f must have a point of period k.*

Proof.   The proposition can now be proved with standard argu-

ments on the appropriate iterates of f. For completeness, we shall give a brief outline. First consider the case when $h = (2m + 1) \cdot 2^s$ and $k = (2n + 1) \cdot 2^t$ where m, n, s, t are nonnegative integers such that either (1) $s < t$ and $m > 0$ or (2) $s = t$ but $0 < m < n$. Applying Proposition 3.1 and Theorem 2.2 respectively in these two cases to the function $g = f^{2^s}$, we may get a periodic point x for g of the period $(2n + 1) \cdot 2^{t-s}$. It is straightforward to show that x is periodic point of period $(2n + 1) \cdot 2^t$ for f.

Then consider the case when $h = (2m + 1) \cdot 2^s$ and $k = 2^t$ for some integer $m \geq 0$ and positive integers s and t such that either (a) $m > 0$, or (b) $m = 0$ but $s \geq t$. Note that in Case (a) we may also assume that $s \geq t$, for otherwise, it reduces to Case (1) of the preceding paragraph. Now with this assumption, apply Proposition 3.2 to the function $g = f^{2^{t-1}}$. In either case we get a point x of period 2 for g. It is straightforward to show that x is of period 2^t for f.

The above two paragraphs together with the implication that any period h implies period $k = 1$ cover all the possibilities for h and k where h precedes k in the Sharkovsky's sequence.

To prove that the Sharkovsky's sequence is strict, we need a kind of converse for Theorem 2.1 so that periodic points can completely be determined by cycles in the digraphs. For this purpose, we introduce the following.

DEFINITION 3.4. Let $f \colon R \to R$ be a continuous function and x_0 be a point of period k for f. The k-periodic digraph of f for the point x_0 is said to be *faithful* if for each vertex I_j in the digraph, $J_1, J_2, \cdots, J_n$ are the vertices to which I_j is directed, then $f(I_j) = \bigcup_{i=1}^{n} J_i$.

PROPOSITION 3.5. *Let $f \colon R \to R$ be a continuous function, x_0 be a point of period k for f, and $I_1, I_2, \cdots, I_{k-1}$ be the finite intervals in the partition of the real line by the images of x_0. Suppose the k-periodic digraph of f for the point x_0 is faithful. Then whenever f has a point of period p in the set $\bigcup_{i=1}^{k-1} I_i$, the k-periodic digraph contains a cycle of length p.*

Proof. Let $b \in I_j$, for some j, be a point of period p for f. Set this interval $I_j = I^0$. Since $f(I_j) = \bigcup_{i=1}^{n} J_i$, there exists an interval I^1 such that $f(b) \in I^1$. Continuing this process, we obtain intervals $I^2, \cdots, I^p$ such that $f(I^i) \supset I^{i+1}$, and for each i, $f^i(b) \in I^i$. Hence, $I^p = I^0$. This yields a p-cycle in the digraph.

REMARK 3.6. The p-cycle obtained in the preceding proposition may be a repetitive cycle in general. However, if we further require

the function f to be monotone over each interval I_i, $i = 1, 2, \cdots$, $k - 1$, then for each $p > 1$, the corresponding p-cycle cannot consist entirely of a 1-cycle traced several times.

PROPOSITION 3.7. *For each $n \geq 2$, there exists a continuous function $f: R \to R$, which has a point of period $2n + 1$, but has no point of period $2m + 1$ for any m with $1 \leq m < n$.*

Proof. Let f be a piecewise linear function whose graph is given by Figure 3.
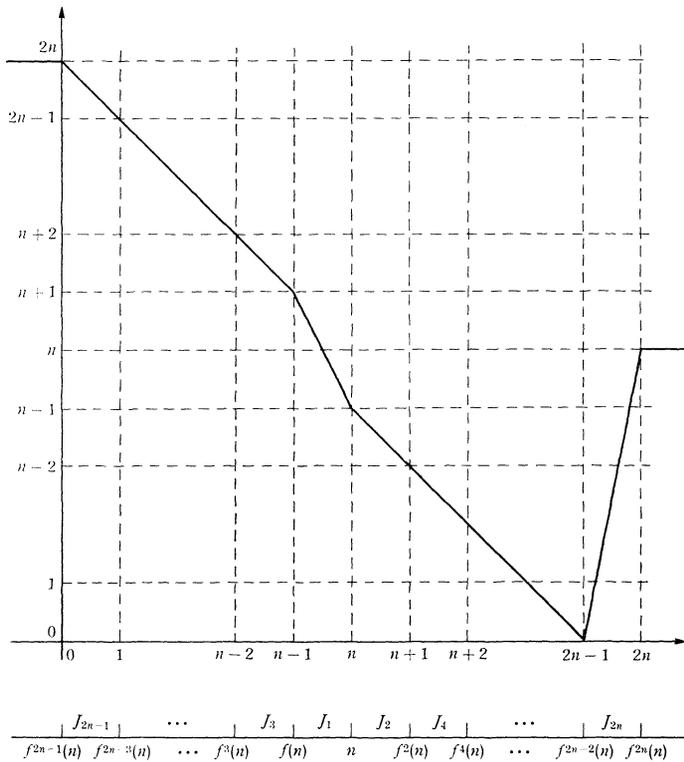
FIGURE 3

The point n is a periodic point of period $2n + 1$ for f. In fact, $f(n) = n - 1$, $f^2(n) = n + 1$, $f^3(n) = n - 2$, $f^4(n) = n+2$, $\cdots$, $f^{2n-1}(n) = 0$, $f^{2n}(n) = 2n$, and $f^{2n+1}(n) = n$. If we let $J_1 = [f(n), n]$, $J_3 = [f^3(n), f(n)]$, $\cdots$, $J_{2n-1} = [f^{2n-1}(n), f^{2n-3}(n)]$ and $J_2 = [n, f^2(n)]$, $J_4 = [f^2(n), f^4(n)]$, $\cdots$, $J_{2n} = [f^{2n-2}(n), f^{2n}(n)]$, then in the corresponding $2n + 1$ digraph, J_1 is directed to J_1 and J_2. For each i, $1 < i < 2n$, the vertex J_i is directed only to J_{i+1}, and J_{2n} is directed to J_{2k-1} for every k with $1 \leq k \leq n$. The digraphs for the cases $n = 3$ and $n = 4$ are given in Figure 4.

Digraph for $n = 3$

Digraph for $n = 4$
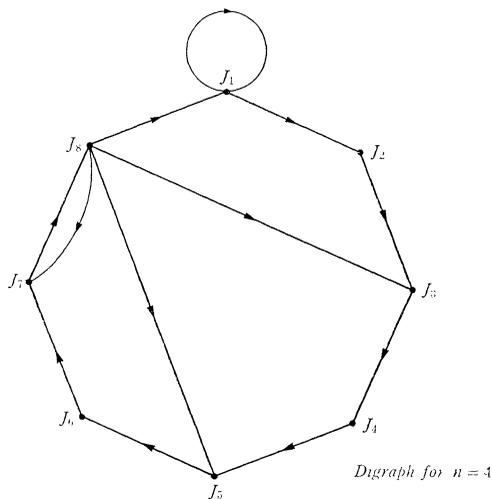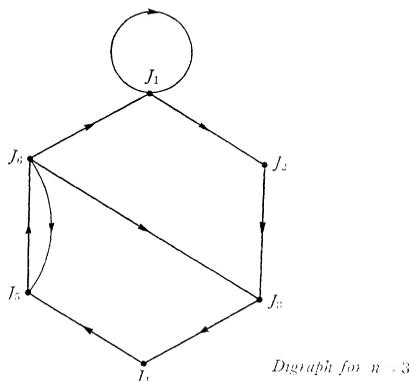
FIGURE 4

Observe that except for a repetition of the 1-cycle at J_1, the only cycles with a length $< 2n + 1$ in the $2n + 1$ periodic digraph are those of an even order. Thus, the proof is completed by Proposition 3.5 and Remark 3.6.

PROPOSITION 3.8. *Let* h *and* k *be two positive integers such that* h *precedes* k *in Sharkovsky's sequence. There always exists a continuous function from* R *into* R *which has a point of period* k *but no point of period* h.

Proof. The proof of this general result follows from the preceding proposition by a more or less standard procedure of taking the "square root" of a function. We shall outline the main idea below. Suppose $f: [a, b] \to [a, b]$ is a continuous function which has a point of period $2n + 1$ in the interval $[a, b]$, but no point of period $2m + 1$

for m with $1 \leq m < n$.   Let $[c, d]$ be an interval disjoint from $[a, b]$, say $b < c$, and let $g: [a, b] \to [c, d]$ be an arbitrary homeomorphism. Then, define a continuous function $r: R \to R$ by letting $r(x) = g \circ f(x)$ for $x \in [a, b]$, $r(x) = g^{-1}(x)$ for $x \in [c, d]$, $r \equiv a$ linear function on the interval $[b, c]$, and $r \equiv$ constant for $(-\infty, a]$ and for $[d, +\infty)$.

Observe that $r^2(x) = f(x)$ for each $x \in [a, b]$ (hence, the name "square root"), and $r^2(x) = gfg^{-1}(x)$ for each $x \in [c, d]$.   But conjugate functions have the same periodic points.   The function r has points of period $(2n + 1) \cdot 2$ but no point of period $(2m + 1) \cdot 2$ for $1 \leq m < n$. Furthermore, r does not have any point of an odd period > 1, for under an odd iterate of r, the points of $[a, b]$ and $[c, d]$ are carried into the "wrong" interval.   Repeated applications of this process yield functions which has points of period $(2n + 1) \cdot 2^t$, but no points of period $(2m + 1) \cdot 2^s$ where either $s < t$, or $s = t$ but $1 \leq m < n$.

It is easy to construct a function which has fixed points in a closed interval but no points of a period > 1.   Repeated applications of the above process on this function will give rise to examples which cover other cases of h and k in the Sharkovsky's sequence.   This completes the proof of Proposition 3.8, and thus, also the proof of Sharkovsky's theorem.

IV.   Some comments.

REMARK 4.1.   Our argument can be modified for functions $f: [a, b] \to [a, b]$ for some interval $[a, b]$, and the examples constructed in Propositions 3.7 and 3.8 can clearly be made into smooth functions. Therefore, one can use Sharkovsky's theorem for different classes of functions, continuous or smooth, defined on a finite interval or on the entire real line.

REMARK 4.2.   Each periodic point for a continuous function $f: R \to R$ gives rise to a periodic digraph.   However, not every directed graph can be the periodic digraph of some function.

Question.   Under what condition can a directed graph be the periodic digraph of some continuous function $f: R \to R$?

REMARK 4.3.   If a directed graph is realizable as a periodic digraph of some function, then there are infinitely many such functions.   For if a directed graph is a k-periodic digraph of some function f for a point x of period k, then any other function g which agrees with f on the k points $x, f(x), \cdots, f^{k-1}(x)$ also has the directed graph as its k-periodic digraph.   However, among these infinitely many functions, the "simplest" ones will be those which are monotone

over each of the intervals in the partition of the real line by the images of x. For by our results in §III, these functions have only those periodic points which correspond to a cycle in the given digraph.

Added in Proof. An independent proof of this theorem, also based on diagraphs, was discovered jointly by Block, Guckenheimer, Misiurewicz and Young. Their paper, "Periodic points and topological entropy of one dimensional maps," will appear in the Proceedings of the Northwestern University Dynamical System Conference (Springer Lecture Notes No. 819).

REFERENCES

1. J. Frauenthal, *Introduction to Population Modeling*, UMAP, Newton, Mass., 1979.
2. J. Guckenheimer, *On the bifurcation of maps of the interval*, Invent. Math., **39** (1977), 165–178.
3. T.-Y. Li and J. A. Yorke, *Period three implies chaos*, Amer. Math. Monthly, **82** (1975), 985–992.
4. A. N. Sharkovsky, *Co-existence of the cycles of a continuous mapping of the line into itself*, Ukranian Math. Z., **16** (1964), 61–71 (Russian).
5. P. Štefan, *A theorem of Šarkovskii on the existence of periodic orbits of continuous endomorphisms of the real line*, Comm. Math. Phys., **54** (1977), 237–248.
6. P. D. Straffin, Jr., *Periodic points of continuous functions*, Math. Mag., **51** (1978), 99–105.

SOUTHERN ILLINOIS UNIVERSITY AT EDWARDSVILLE
EDWARDSVILLE, IL 62026