From Intermediate Value Theorem To Chaos

XUN-CHENG HUANG
New Jersey Institute of Technology
Newark, NJ 07102

1. Introduction

Continuous functions of a single variable have been studied extensively for over 200 years. Great mathematicians such as Newton (1642–1727), Leibniz (1646–1716), and Euler (1707–1783) have left enduring monuments in this field. Their rich achievements (for example, Euler published 886 papers and books in his 76 years of life) now comprise the major part of the calculus with which every student of science and engineering is familiar. It is hard to believe that in such a field, ploughed and cultivated repeatedly by so many great masters, there is still some virgin land.

In 1975, the article: "Period three implies chaos", was published in the *American Mathematical Monthly* by Li and Yorke. (Here, the word "period" is used in a different way from that seen in elementary mathematics. For example, period three means that there is a point x_0 such that $f^3(x_0) = f(f(f(x_0))) = x_0$, $f^k(x_0) \neq x_0$ for $k = 1, 2$; in other words, the image of x_0 comes back to x_0 after three iterations.) In that article, Li and Yorke announced that a new theorem for continuous functions of a single variable was discovered. The theorem states that if a continuous function has period three, it must have period n for every positive integer n. Soon afterwards, it was found that Li and Yorke's theorem is only a special case of a remarkable theorem published a decade earlier by Soviet mathematician A. N. Sarkovskii, in a Ukrainian journal. Sarkovskii reordered the natural numbers and proved that if $l \lhd m$ (which means l is "less than" m in Sarkovskii's ordering) and if a function has period l then it must have period m. The number 3 is the "smallest" in Sarkovskii's ordering. So, obviously, period 3 implies all the other periods, and Li-Yorke's theorem was not a new one. However, it was in Li-Yorke's article that the new concept of chaos was first introduced into mathematics. People were surprised that iterations of even a very simple continuous function of a single variable can display extremely complicated chaotic behavior.

The original proof of Sarkovskii's theorem is quite difficult. More recently, several authors have simplified the proof (see for example, [3, 5]), however, their proofs are still overly complicated. In this article, we introduce a proof that is based on the intermediate value theorem, accessible to readers with some knowledge of calculus.

2. Common Facts and the Intermediate Value Theorem

Without any special knowledge of mathematics, one can understand the following common facts:

Two trains that depart at the same time from New York and Chicago, destined for Chicago and New York, respectively, must meet each other along their trips.

In a marathon race, a contestant who is lagging behind at first and wants to win must catch up and pass all the other contestants.

Now, let's play a little trick on these common facts. Is it still so obvious?

Suppose Robert starts to climb up a mountain at 8 a.m. and reaches the top at 6 p.m., and then at the same time next day begins his return using the same route. Is there any place on his way up and down the mountain where his watch indicates the same time?

The answer is yes. We can imagine two Roberts start at the same time, one climbing up and the other climbing down by the same route. If their watches are adjusted before starting, of course, they will show exactly the same time where the two Roberts meet on their ways.

The above idea can be summarized in mathematics as the following theorem:

INTERMEDIATE VALUE THEOREM. *If f is continuous on $[a,b]$ and N is any number between $f(a)$ and $f(b)$, then there exists at least one x_0 between a and b such that $f(x_0)=N$.*

This is one of the most fundamental theorems in calculus. Although it is very simple and ordinary, people still pay great attention to it and use it as a test question. For example, the following question is often asked: Prove

PROPOSITION 2.1. *Let f be continuous on $[a,b]$. If the range of f contains $[a,b]$, then equation*

$$f(x)=x$$

has at least one solution in $[a,b]$.

The solution is straightforward. Since the range of f contains $[a,b]$, there must be some $x_1, x_2 \in [a,b]$ such that (see FIGURE 1)

$$f(x_1) \leqslant a \leqslant x_1, \qquad f(x_2) \geqslant b \geqslant x_2.$$

Let $g(x)=f(x)-x$. The result follows from applying the intermediate value theorem with $N=0$.

By the way, if the assumption "the range of f contains $[a,b]$" is replaced by "the range of f is contained in $[a,b]$," Proposition 2.1 is still valid (see FIGURE 2).

A point x_0 satisfying equation (2.1) is called a *fixed point*. A natural generalization of fixed point is *periodic point*.

Assume $R\{f\} \subset D\{f\}$, the range of f is contained in the domain of f. Denote $f^0(x)=x$, $f^1(x)=f(x)$, $f^2(x)=f(f(x))$, $f^3(x)=f(f^2(x)),\dots,f^n(x)=f(f^{n-1}(x))$. If x_0 satisfies

$$\begin{cases} f^n(x_0)=x_0 \\ f^k(x_0)\neq x_0, \qquad k=1,2,\dots,n-1, \end{cases} \tag{2.1}$$

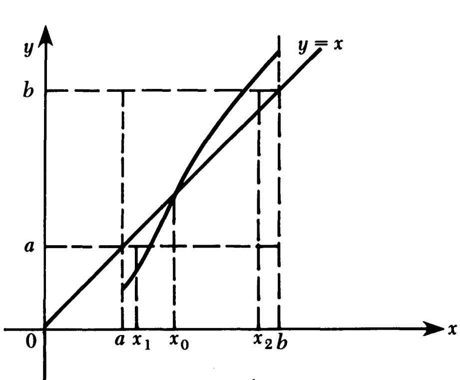


FIGURE 1

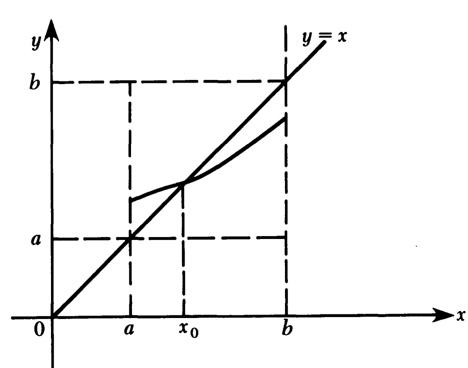


FIGURE 2

then x_0 is called an *n-periodic point with period n*. Clearly, a fixed point is a 1-periodic point.

If x_0 is an n-periodic point of f, then $x_0, f(x_0), \ldots, f^{n-1}(x_0)$ are distinct and the set $\{x_0, f(x_0), \ldots, f^{n-1}(x_0)\}$ is called a *periodic orbit* of f.

If f has an n-periodic point, we say that *f has period n*.

The existence of a fixed point of a function is generally clear by the inspection of its graph. But the existence of an n-periodic point is not so easy to see even if n is a small integer. As an example, let us consider the function

$$\psi(x) = \begin{cases} x + \frac{1}{2}, & 0 \leqslant x \leqslant \frac{1}{2} \\ 2 - 2x, & \frac{1}{2} < x \leqslant 1, \end{cases} \tag{2.2}$$

whose graph is in FIGURE 3.

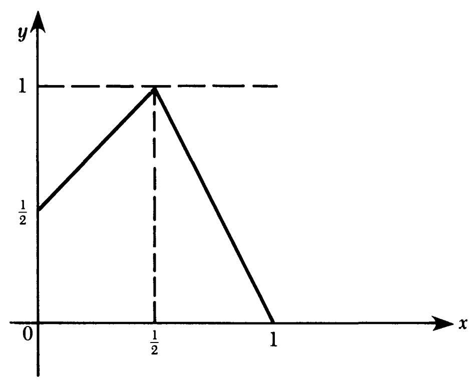


FIGURE 3

As you can see, $\psi(0) = \frac{1}{2}$, $\psi^2(0) = \psi(\frac{1}{2}) = 1$, $\psi^3(0) = \psi^2(\frac{1}{2}) = \psi(1) = 0$; that is ψ has a 3-periodic point 0. Does it have a 5-periodic point? A 7-periodic point? It is hard to ascertain this by just looking at the graph. We need to do some deeper analysis.

The following is a generalized version of the Intermediate Value Theorem.

PROPOSITION 2.2. *Let f be continuous on $[a, b]$, and let $I_0, I_1, \ldots, I_{n-1}$ be closed subintervals of $[a, b]$. If*

$$f(I_k) \supset I_{k+1}, \qquad k = 0, 1, \ldots, n-2,$$
$$f(I_{n-1}) \supset I_0, \tag{2.3}$$

then, the equation

$$f^n(x) = x \tag{2.4}$$

has at least one solution $x = x_0 \in I_0$ such that

$$f^k(x_0) \in I_k, \qquad k = 0, 1, \ldots, n-1. \tag{2.5}$$

In the proposition, $f(I_k) \supset I_{k+1}$ means that the range of f on I_k contains I_{k+1}. We will use the notation

$$I_i \to I_j \qquad \text{or} \qquad I_j \leftarrow I_i \tag{2.6}$$

if $f(I_i) \supset I_j$ ($f(I_i)$ "covers" I_j). The condition (2.3) can be written as

$$I_0 \to I_1 \to I_2 \to \cdots \to I_{n-1} \to I_0.$$

Clearly, if $n = 1$, Proposition 2.2 is reduced to Proposition 2.1.

The proof of Proposition 2.2 is based on the following fact:

If $I_1 \to I_2$, then there exists a subinterval $I_1^* \subset I_1$ such that $f(I_1^*) = I_2$.

This is true, since if $I_2 = [c, d]$, there exist x_1 and x_2 in I_1 such that $f(x_1) = c$ and $f(x_2) = d$. Let $I_1^* = [x_1, x_2]$. Then $I_1^* \subset I_1$, and by the intermediate value theorem, $f(I_1^*) = I_2$.

This fact implies that there exist $I_{n-1}^* \subset I_{n-1}$ such that $f(I_{n-1}^*) = I_0$, $I_{n-2}^* \subset I_{n-2}$ such that $f(I_{n-2}^*) = I_{n-1}^*, \dots,$ and $I_0^* \subset I_0$ such that $f(I_0^*) = I_1^*$. In other words, there exist $I_k^* \subset I_k$ such that

$$\begin{aligned} f(I_k^*) &= I_{k+1}^* \subset I_{k+1} \quad \text{for } k = 0, 1, \dots, n-2 \\ \text{and} \qquad f(I_{n-1}^*) &= I_0 \supset I_0^*. \end{aligned} \tag{2.7}$$

From (2.7), $f^k(I_0^*) = I_k^*$, for $k = 0, 1, \dots, n-2$, and $f^n(I_0^*) \supset I_0^*$. Thus, by Proposition 2.1, equation (2.4) has a solution $x_0 \in I_0^* \subset I_0$ such that (2.5) holds.

Note: In geometry, (2.5) means that mapped successively by f, x_0 visits $I_1, I_2, \dots, I_{n-1}$ and finally comes back to where it was.

Proposition 2.2 itself is not a remarkable result. But it is the only calculus we need for the proof of our main result.

Proposition 2.3. *Let f: $I \to I$ be continuous, and let f have a $(2n+1)$-periodic orbit $\{x_k = f^k(x_0), k = 0, 1, \dots, 2n\}$, but no $(2m+1)$-periodic orbit for $1 \leqslant m < n$. Suppose x_0 is in the middle of all the x_i's. Then one of the two permutations*

$$\begin{aligned} &\text{(i)}\ \ x_{2n} < x_{2n-2} < \cdots < x_2 < x_0 < x_1 < \cdots < x_{2n-3} < x_{2n-1} \\ &\text{(ii)}\ \ x_{2n-1} < x_{2n-3} < \cdots < x_1 < x_0 < x_2 < \cdots < x_{2n-2} < x_{2n} \end{aligned} \tag{2.8}$$

is valid. (Figure 4 *is for the case* $n = 3$).

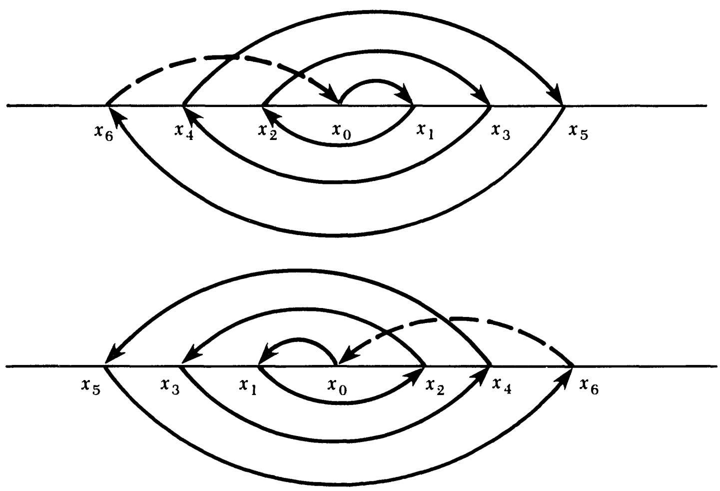


FIGURE 4

Proof. Suppose $n > 1$. Reorder $\{x_i, i = 0, 1, \dots, 2n\}$ as $\{z_i, i = 1, 2, \dots, 2n+1\}$ such that

$$z_1 < z_2 < \cdots < z_{2n+1}.$$

Let S_{kl} be the set $\{z_i, k \leqslant i \leqslant l\}$. Assume

$$\min\{f(z): z \in S_{kl}\} = z_i$$
$$\max\{f(z): z \in S_{kl}\} = z_j,$$

and define the set function f^* as

$$f^*(S_{kl}) = S_{ij}. \tag{2.9}$$

We use the notation $S_{kl} \to S_{ij}$, to denote $f^*(S_{kl}) \supset S_{ij}$. Our proof is based on the following claim:
There exist

(i) positive integers m, l $(m, l < 2n+1)$,

(ii) a family of sets $S_i = S_{k_i l_i}$ $(i = 1, 2, \dots, 2n)$ such that

$$S_1 = \{z_m, z_{m+1}\}, \qquad S_{2n} = \{z_l, z_{l+1}\},$$

S_i has only one common point with S_{i+1}, and

$$\begin{gathered} S_1 \to S_2 \to \cdots \to S_{2n} \to S_1 \\ S_1 \subset S_2 \subset \cdots \subset S_{2n-1} \not\supset S_{2n}. \end{gathered} \tag{2.10}$$

In fact, since $f(z_1) > z_1$ and $f(z_{2n+1}) < z_{2n+1}$, we can choose the largest i, say m, such that $f(z_m) > z_m$. Clearly, $m \leqslant 2n$.

Let $S_1 = \{z_m, z_{m+1}\}$, $S_2 = f^*(S_1), \dots, S_{i+1} = f^*(S_i), \dots$. Since x_0 is not a 2-periodic point, we have

$$S_{i+1} \supset S_i,\ i = 1, 2, 3, \dots, \quad \text{and}$$
$$S_i \not\supset S_{i+1} \text{ if } S_i \text{ is not the set } \{z_1, z_2, \dots, z_{2n+1}\}.$$

Suppose that $i = 1, 2, \dots, t-1$. If $t = 2n$, then (2.10) is valid. We only need to prove that $t = 2n$.

Since the number of points in $S_{1m} = \{z_1, \dots, z_m\}$ differs from that in $S_{m+1,2n+1} = \{z_{m+1}, \dots, z_{2n+1}\}$, there exists $l \neq m$ such that $f(z_l)$ and $f(z_{l+1})$ are on different sides of $[z_m, z_{m+1}]$. Thus, $[z_l, z_{l+1}] \to [z_m, z_{m+1}]$.

Let S_t be $\{z_l, z_{l+1}\}$. Here $t-1$ is chosen as the smallest i such that $S_i \to \{z_l, z_{l+1}\}$. This is possible because $S_1 \subset S_2 \subset S_3 \subset \dots$.

As an illustration, in the first case of FIGURE 5, $S_1 = \{z_m, z_{m+1}\} = \{z_4, z_5\}$, $S_2 = \{z_3, z_5\}$, $S_3 = \{z_3, z_6\}$, $S_4 = \{z_2, z_6\}$, $S_5 = \{z_2, z_7\}$, $S_t = \{z_l, z_{l+1}\} = \{z_1, z_2\}$, $S_{1m} = \{z_1, z_2, z_3, z_4\}$, and $S_{m+1,2n+1} = \{z_5, z_6, z_7\}$.

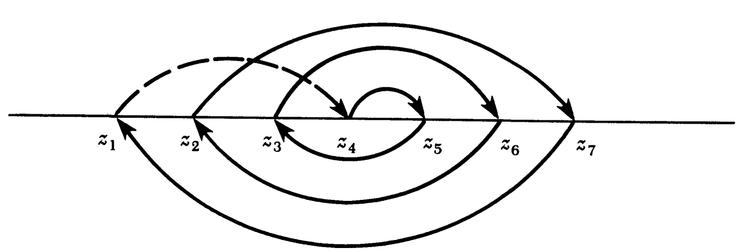


FIGURE 5

Assume that I_i is the smallest closed interval that contains S_i. Then

$$\begin{aligned} &I_1 \subset I_2 \subset \cdots \subset I_{t-1} \not\supset I_t \\ &I_1 \to I_2 \to \cdots \to I_t \to I_1. \end{aligned} \tag{2.11}$$

Since $S_{t-1} \not\supset S_t = \{z_l, z_{l+1}\}$, and S_{t-1} contains at least t points, we have $t \leqslant 2n$.

Assume that $t < 2n$. Let $J_0 = J_1 = \cdots J_{2n-t-1} = I_1$, $J_{2n-t} = I_2$, $J_{2n-t+1} = I_3, \ldots$ $J_{2n-2} = I_t$. We have

$$J_0 \to J_1 \to J_2 \to \cdots J_{2n-2} \to J_0. \tag{2.12}$$

By Proposition 2.2, there exists $x_0^* \in J_1$, such that

$$\begin{aligned} &f^{2n-1}(x_0^*) = x_0^* \\ \text{and}\qquad &f^k(x_0^*) \in J_k \quad \text{for } k = 0, 1, \ldots, 2n-2. \end{aligned} \tag{2.13}$$

It is easy to see that $x_0^*, f(x_0^*), \ldots, f^{2n-2}(x_0^*)$ are distinct. For otherwise, the period of x_0^* is less than $2n-1$, and consequently $f^{2n-2}(x_0^*)$ would then be equal to one of $x_0^*, f(x_0^*), \ldots, f^{2n-3}(x_0^*)$. Then by (2.13)

$$f^{2n-2}(x_0^*) \in J_0 \cap J_{2n-2} = I_1 \cap I_t. \tag{2.14}$$

That is impossible because, by the construction of I_i, $I_1 \cap I_t$ is empty for $t > 2$. Thus, $t = 2n$, and $S_{i+1} \setminus S_i$ is a singleton for each $i = 1, 2, \ldots 2n-2$.

For Proposition 2.3, it is sufficient to show that for each $i = 1, 2, \ldots 2n-1$, f always maps one end point of S_i, say A, to the other, say B, and A is always between B and $f(B)$. Let $S_i \setminus S_{i-1} = \{A_i\}$. If it is not the case for some $k < 2n$, then we have

$$\curvearrowleft [A_{k-2}, A_k] \leftrightarrows [A_{k-1}, A_{k-3}] \tag{2.15}$$

(See FIGURE 6).

Clearly, these two cases will lead to the fact that f has period 3 by Proposition 2.2.

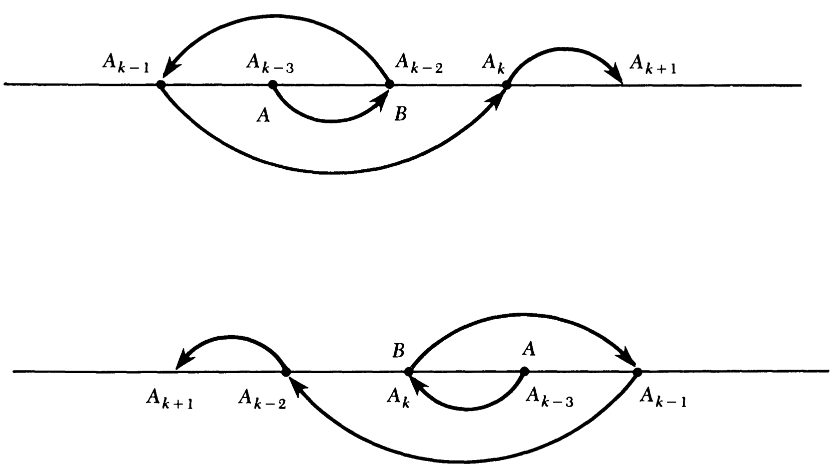


FIGURE 6

3. Period Three and Chaos

The famous Li-Yorke theorem is the following:

THEOREM 3.1. *Let f be continuous on $[a,b]$, its range contained in $[a,b]$. If f has a 3-periodic point, then f has n-periodic points for all positive integers n.*

Proof. Let $x_0 < x_1 < x_2$ be a 3-periodic orbit of f. Then either $f(x_1) = x_0$ or $f(x_1) = x_2$. Without loss of generality, suppose $f(x_1) = x_0$. Then $f(x_0) = x_2$, $f(x_2) = x_1$. Let $\tilde{I}_0 = [x_0, x_1]$, $\tilde{I}_1 = [x_1, x_2]$. By the intermediate value theorem

$$\circlearrowleft \tilde{I}_0 \leftrightarrows \tilde{I}_1. \tag{3.1}$$

Let

$$\begin{aligned} I_0 = I_1 = \cdots I_{n-2} &= \tilde{I}_0 \\ I_{n-1} &= \tilde{I}_1. \end{aligned} \tag{3.2}$$

Proposition 2.2 implies there exists $x_0^* \in \tilde{I}_0$ such that $f^n(x_0^*) = x_0^*$ and

$$\begin{aligned} f^k(x_0^*) &\in \tilde{I}_0, \qquad k = 0, 1, \dots, n-2 \\ f^{n-1}(x_0^*) &\in \tilde{I}_1. \end{aligned} \tag{3.3}$$

By the same argument as for (2.14), if $x_0^*, f(x_0^*), \dots, f^{n-1}(x_0^*)$ is not an n-periodic point of f, then $f^{n-1}(x_0^*)$ would be one of $f^k(x_0^*)$, $k = 0, 1, \dots, n-2$. Thus,

$$f^{n-1}(x_0^*) \in \tilde{I}_0 \cap \tilde{I}_1 = x_1,$$

and

$$x_0^* = f^n(x_0^*) = x_0$$

$$f(x_0^*) = f(x_0) = x_2 \notin \tilde{I}_0,$$

which is a contradiction to $f(x_0^*) \in \tilde{I}_0$.

The proof of Theorem 3.1 is completed.

Theorem 3.1 tells us that the function whose graph is shown in FIGURE 3 has period n for each n. It is beyond one's imagination that such a simple function is such a complicated phenomenon.

In [2] the new concept "chaos" was first introduced. The meaning of chaos in mathematics is that if f has a 3-periodic point in I then there exists an uncountable set $S \subset I$ such that for any two points $x_0, y_0 \in S$, the distance between the two iterative series $x_n = f^n(x_0)$, $y_n = f^n(y_0)$, $n = 1, 2 \dots$, has the property that, as $n \to \infty$, the limit inferior equals zero while the limit superior is greater than zero.

Clearly the points in S have very interesting properties under successive mapping by f. That the limit inferior equals zero means that there are infinitely many n such that $\{f^n(x)\}$ and $\{f^n(y)\}$ are as close as you like, and that the limit superior is greater than zero means that there are infinitely many n such that the distance between $\{f^n(x)\}$ and $\{f^n(y)\}$ is always positive. In other words, under the successive iteration of f, different points of S are sometimes close, sometimes separated, and none of them is periodic.

4. Sarkovskii's Theorem

The order of natural numbers is

$$1,2,3,4,5,\dots.$$

But not many people know that they can also be reordered as the following:

$$3,5,7,\dots,2\cdot 3,2\cdot 5,2\cdot 7,\dots,2^2\cdot 3,2^2\cdot 5,2^2\cdot 7,\dots,16,8,4,2,1. \qquad (4.1)$$

That is, first list all odd numbers except 1, followed by 2 times the odds, 2^2 times the odds, 2^3 times the odds, etc. This exhausts all the natural numbers with the exception of the powers of two that are listed at the end in decreasing order. The number 1 is last.

The ordering (4.1) is now known as *Sarkovskii's ordering* of natural numbers, and is denoted as

$$\begin{gathered} 3\lhd 5\lhd 7\lhd \cdots \lhd 2\cdot 3\lhd 2\cdot 5\lhd 2\cdot 7\lhd \cdots \lhd \\ 2^2\cdot 3\lhd 2^2\cdot 5\lhd 2^2\cdot 7\lhd \cdots \lhd 16\lhd 8\lhd 4\lhd 2\lhd 1. \end{gathered} \qquad (4.2)$$

Is there any application of Sarkovskii's ordering? Let us consider the following theorem.

THEOREM 4.1 (Sarkovskii, 1964). *Let $f\colon I\to I$ be continuous and let f have an l-periodic point. If $l\lhd m$, then f has an m-periodic point, too.*

Here, I can be any interval, finite or infinite, open or closed, semi-open or semi-closed. This theorem tells us that the periods of a continuous function show a wonderful regularity. Before proving the theorem, we make several remarks:

1. If f has a periodic point whose period is not a power of two, then f must have infinitely many periodic points. Conversely, if f has only finitely many periodic points, then each period must be a power of two. Here, the number 1 is considered as 2^0.
2. Period 3 is the least period in the Sarkovskii ordering and therefore implies the existence of all other periods as we saw in Theorem 3.1.
3. The converse of Sarkovskii's Theorem is also true. There exist functions that have p-periodic points and no "higher" periodic points in the sense of Sarkovskii's ordering.

Proof of Theorem 4.1.

Case 1. If f has period 2^m, then f has period 2^l for each $l<m$.

We just need to prove that period 2^m implies period 2^{m-1}.

If $m=1$, then f has a 2-periodic point. Let x_1,x_2 $(x_1<x_2)$ be a 2-periodic orbit of f. That is, $f(x_1)=x_2$, $f(x_2)=x_1$, or $f([x_1,x_2])\supset[x_1,x_2]$. By Proposition 2.1, f has a fixed point (i.e. 2^0-periodic point).

Suppose the conclusion is valid for $m=k$. We want to show it is also valid for $m=k+1$.

Let $g=f^2$. Then f has period 2^{k+1} implies that g has period 2^k. Now by the induction hypothesis, g has period 2^{k-1}. That is, there exists an $x_0\in I$ such that

$$g^{2^{k-1}}(x_0)=x_0.$$

$$g^t(x_0) \neq x_0 \quad \text{for } t = 1, 2, \dots, 2^{k-1},$$

which is equivalent to

$$f^{2^k}(x_0) = x_0,$$
$$f^{2t}(x_0) \neq x_0 \quad \text{for } t = 1, 2, \dots, 2^{k-1}.$$

Suppose that x_0 is not a 2^k-periodic point of f. Then there must be some $s \in \{1, 3, 5, \dots, 2^k - 1\}$ such that

$$f^s(x_0) = x_0.$$

But, it is impossible because this implies that

$$f^{2s_0}(x_0) = x_0$$

for some $s_0 \in \{1, 2, 3, \dots, 2^{k-1} - 1\}$.

Therefore, the induction completes the proof.

Case 2. If f has period $2m + 1$, $(m > 1)$, then f has period k for all $k > 2m + 1$.

By Proposition 2.3, letting $I_1 = [x_0, x_1]$, $I_2 = [x_2, x_0], \dots, I_{2n-1} = [x_{2n-3}, x_{2n-1}]$, $I_{2n} = [x_{2n}, x_{2n-2}]$, we have the following diagram (FIGURE 7)

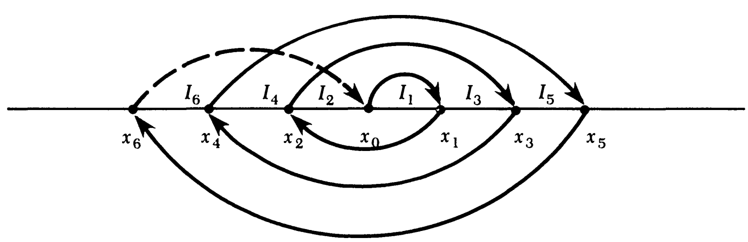
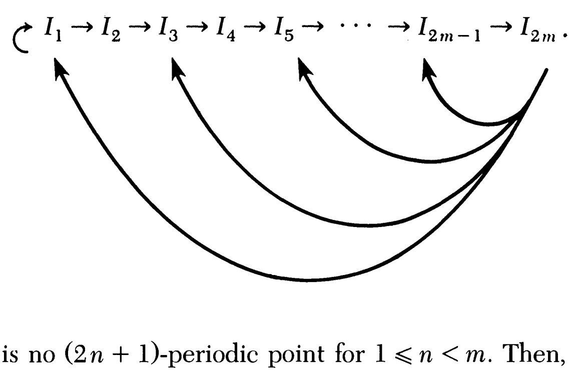


FIGURE 7

and



(4.3)

Suppose there is no $(2n + 1)$-periodic point for $1 \leqslant n < m$. Then, for $k > 2m + 1$

$$\underbrace{I_1 \to I_1 \to \cdots \to I_1}_{k-(2m-1)} \to I_2 \to \cdots I_{2m} \to I_1. \tag{4.4}$$

Since there is no common point in I_1 and I_{2n-1} Proposition 2.2 will result in f having a k-periodic point.

Case 3. If f has period $2m + 1$, $m > 1$, then f has period $2k$ for any positive integer k.

We only need to prove the case where $2k \leqslant 2m$. From (4.3), we have

$$I_{2(m-k)+1} \to I_{2(m-k)+2} \to \cdots I_{2m} \to I_{2(m-k)+1}.$$

By the same argument as in case 2, f has a $2k$-periodic point.

Case 4. Let $m \lhd n$, where $m = 2^k p$, $n = 2^t q$, p, q are odd numbers, $p > 1$, $k \geqslant 1$. Then period m implies period n. Without loss of generality, suppose that for $l \lhd m$ there is no l-periodic point of f.

According to the Sarkovskii ordering, we need to consider the following possibilities:

$$\text{(i) } t > k,\ q \geqslant 1, \quad \text{and}$$
$$\text{(ii) } t = k,\ q > p.$$

Let $g(x) = f^{2^k}(x)$. Then f has period $2^k p$ implies g has period p.

By Case 3, g has period $2^{t-k}q$ for $t > k$ and $q \geqslant 1$. Therefore f has period $2^t q$ for $t > k$, $q \geqslant 1$, and (i) is valid.

By Case 2, g having period p implies that g has period q. Then f has period $2^t q$, and (ii) is valid.

The proof of Theorem 4.1 is complete.

The following example shows that period 5 does not imply period 3.

Let f be the piecewise linear function defined on $[1,5]$ with $f(1) = 3$, $f(2) = 5$, $f(3) = 4$, $f(4) = 2$ and $f(5) = 1$, whose graph is shown in FIGURE 8.

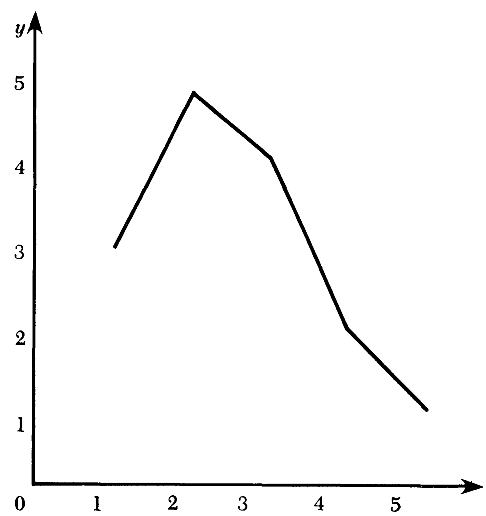


FIGURE 8

It is easy to check that

(i) $10/3$ is a fixed (or 1-periodic) point;

(ii) $5/3$ is a 2-periodic point;

(iii) $1,2,3,4,5$ are 5-periodic points.

We can also prove that f has no 3-periodic point. Since

$$f^3[1,2] = [2,5],\ f^3[2,3] = [3,5],\ f^3[4,5] = [1,4],$$

f^3 has no periodic point in any of these intervals. Also, since $f^3[3,4] = [1,5]$ and f^3 is monotonically decreasing on $[3,4]$, there exists a unique $x_0 \in [3,4]$ such that

$$f^3(x_0) = x_0.$$

Since $f(x) = 10 - 2x$ on [3,4], $f(x)$ has a unique fixed point $\bar{x} = 10/3$ on [3,4]. Since,

$$f^3(\bar{x}) = f^2(\bar{x}) = f(\bar{x}) = \bar{x} = x_0,$$

x_0 is not a 3-periodic point. Hence, f has no 3-periodic point.

5. Conclusion and Discussion

Sarkovskii's theorem has not ended the discussion of the periodic orbits of continuous functions. On the contrary, it created a new direction for studying the problem. Many articles and books have been appearing. The question is why so many great classical analysts didn't discover such an important theorem. The reason is that the classical analysts concentrated on the local properties of functions. Essentially, continuity, differentiability, and integrability are determined by the local properties of functions. Although some global properties such as uniform continuity had been obtained, these global properties can be derived simply from the local properties.

A remarkable advance in modern analysis is viewing the situation as a whole in the study of functions. Actually, the concepts such as iteration and periodic orbits have inseparable relations to the global properties. For example, $f(x)$ can be iterated on $[a, b]$ but may not be iterated on any subinterval of $[a, b]$.

Studying the global properties of functions or mappings now forms a new branch of mathematics called global analysis, which includes differential dynamical systems, global differential geometry, qualitative theory of differential equations, differential topology, etc. It is one of the main directions in modern mathematics.

Now we use another example to show that even in fundamental problems one can benefit from taking into account the global structure.

Example. Monkeys' Apples

There was a pile of apples on the beach that belonged to five monkeys and was to be distributed equally among them. The first monkey came and waited for a while but no others followed. He divided those apples into five piles each of which had the same number of apples. But one was left and he threw it into the sea and went away with his own pile of apples. The second monkey came and divided the rest of the apples into five piles equally, too. Again, one was left and was thrown into the sea. Then he went away with his own apples, too. Later, one by one, each monkey did the same as the first two did.

What is the least number of apples on the beach in the beginning? What is the least number of apples left after all the monkeys take away theirs?

The problem is not easy to solve if you use the usual equations. So the famous physicist Dirac suggested doing it as follows.

Let N be the number of apples in the beginning, and A_1, A_2, A_3, A_4, A_5 be numbers of apples taken by the monkeys. Then, we will have a system of linear equations

$$\begin{cases} N - 5A_1 = 1 \\ 4A_1 - 5A_2 = 1 \\ 4A_2 - 5A_3 = 1 \\ 4A_3 - 5A_4 = 1 \\ 4A_4 - 5A_5 = 1 \end{cases}, \tag{5.1}$$

which possesses a particular solution

$$(N, A_1, A_2, A_3, A_4, A_5) = (-4, -1, -1, -1, -1, -1). \tag{5.2}$$

The corresponding homogenous equations of (5.1) have a general solution

$$\left(5\left(\frac{5}{4}\right)^4 k, \left(\frac{5}{4}\right)^4 k, \left(\frac{5}{4}\right)^3 k, \left(\frac{5}{4}\right)^2 k, \frac{5}{4}k, k\right) \tag{5.3}$$

where k is any constant. Therefore, the general solution of (5.1) is

$$\left(5\left(\frac{5}{4}\right)^4 k - 4, \left(\frac{5}{4}\right)^4 k - 1, \left(\frac{5}{4}\right)^3 k - 1, \left(\frac{5}{4}\right)^2 k - 1, \frac{5}{4}k - 1, k - 1\right). \tag{5.4}$$

From (5.4), we can determine that the least positive integer for N is $5^5 - 4 = 3121$ when $k = 4^4 = 256$; and the number of apples left is $4A_5 = 4(k-1) = 1020$.

As you can see this solution is based on the structure of solutions of linear equations. If you don't know the theory, it is really hard to find it.

The method we used for this problem is fundamental and very simple. Suppose x is the number of apples before a monkey came and y the number after he left. Clearly, y is determined by x, say $y = f(x)$, and

$$f(x) = \frac{4}{5}(x-1). \tag{5.5}$$

If there were N apples at first and M apples at last, then

$$M = f(f(f(f(f(N))))) = f^5(N). \tag{5.6}$$

Now, we consider how to get a formula for $f^5(x)$. We rewrite $f(x)$ as

$$f(x) = \frac{4}{5}(x+4) - 4 \tag{5.7}$$

where, as you can see, -4 is a fixed point of $f(x)$.

Obviously,

$$\begin{aligned} f^2(x) &= \left(\frac{4}{5}\right)^2 (x+4) - 4 \\ f^3(x) &= \left(\frac{4}{5}\right)^3 (x+4) - 4 \\ f^4(x) &= \left(\frac{4}{5}\right)^4 (x+4) - 4 \\ f^5(x) &= \left(\frac{4}{5}\right)^5 (x+4) - 4 \end{aligned} \tag{5.8}$$

and hence

$$M = \left(\frac{4}{5}\right)^5 (N+4) - 4. \tag{5.9}$$

In order to have a positive integer M, $N + 4$ must be a multiple of 5^5. So the least positive integer value of N is

$$N = 5^5 - 4 = 3121,$$

and consequently

$$M = 4^5 - 4 = 1020.$$

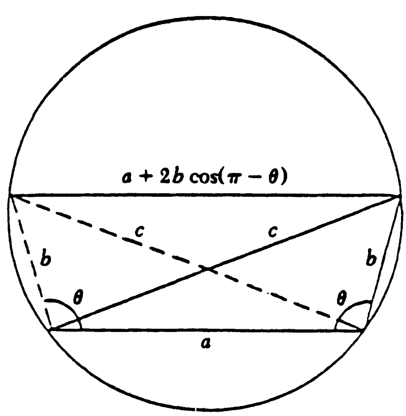
Before ending the article, we mention again that the proof of Sarkovskii's beautiful theorem is far from advanced mathematics. This big surprise shows that people need not have advanced knowledge to establish mathematics if, when opportunity arrives, it is recognized.

Acknowledgements. The author would like to thank Professors Stephen Merrill and Roman Voronka for their useful comments.

REFERENCES

1. A. N. Sarkovskii, Coexistence of cycles of a continuous map of a line into itself, *Ukr. Math. Z.* 16 (1964), 61–71.
2. T. Li and J. A. Yorke, Period three implies chaos, *Amer. Math. Monthly* 82 (1975), 985–992.
3. P. Stefen, A theorem of Sarkovskii on the existence of periodic orbits of continuous endomorphisms of the real line, *Commun. Math. Phys.* 54 (1977), 237–248.
4. Z. Jing, From ordinary fact to surprising theorem, *Nature J.* 8–7 (1985), 532–536.
5. R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Benjamin/Cummings Publishing Co., Menlo Park, CA, 1986.
6. M. Liang, Tidbits of chaos theory and homotopy method, *Nature J.* 9:2 (1986), 139–142.

Proof without Words: The law of cosines via Ptolemy's Theorem



$$c \cdot c = b \cdot b + (a + 2b\cos(\pi - \theta)) \cdot a$$
$$c^2 = a^2 + b^2 - 2ab \cdot \cos\theta$$

—SIDNEY H. KUNG
JACKSONVILLE UNIVERSITY
JACKSONVILLE, FL 32211