

## Generalized local test for local extrema in single-variable functions

Eleftherios Gkioulekas\*

*University of Texas-Pan American, Edinburg, TX 78539-2999*

*(Received 18 Nov 2012, Accepted 22 Jan 2013, Published online: 07 May 2013)*

We give a detailed derivation of a generalization of the second derivative test of single-variable calculus which can classify critical points as local minima or local maxima (or neither), whenever the traditional second derivative test fails, by considering the values of higher-order derivatives evaluated at the critical points. The enhanced test is local, in the sense that it is only necessary to evaluate all relevant derivatives at the critical point itself, and it is reasonably robust. We illustrate an application of the generalized test on a trigonometric function where the second derivative test fails to classify some of the critical points.

### 1. Introduction

As is well-known, the second derivative test is part of the standard single-variable calculus curriculum, and can be used to classify most critical points that are candidates for local minimum or local maximum with zero first derivative. As such, it has some disadvantages. First, it cannot classify critical points where the function loses differentiability or points that are not interior to the domain of the function that is being optimized. According to the Fermat theorem, these are also potential candidates for local minima and maxima. Second, for some functions, finding and factoring the first derivative can be a lot of work already, and finding the second derivative can require an unreasonable amount of effort. All of this is compounded by the third reason, which is that after long calculations to find the second derivative, one might discover that for some critical point the second derivative test is inconclusive.

For these reasons, I have always recommended to students that for most functions, they should just use a table of signs to determine the sign of the first derivative over all relevant intervals, and rely instead on the first derivative test. Unlike the second derivative test, a sign table of the first derivative can always classify all critical points. On the other hand, the second derivative test, due to its local nature, can be especially advantageous for trigonometric functions where, due to periodicity, applying the first derivative test is too awkward. It can also turn up in more advanced coursework in the context of proving theorems that are dependant on the second derivative test. Furthermore, an interesting multidimensional generalization of the second derivative test [1] is essential for rendering mathematically rigorous solutions to constrained optimization problems via the method of Lagrange multipliers.

Obviously, in all of these applications, the risk remains that for specific problems, the second derivative test can turn out to be inconclusive. For multidimensional problems, it is possible to fall back to a generalized first derivative test [2], which is not a well-known result and not part of standard calculus curricula. For single-

---

\*Corresponding author. Email: drlf@hushmail.com

variable problems, one may use a generalization of the second derivative test by Wu [3], where the sign of the second derivative in a local neighborhood around the candidate point is used to classify it as local minimum or maximum or neither. Wu's generalization has the advantage that it is never inconclusive and can salvage the calculation of the second derivative by putting it to good use. As such, it fits with a methodology where one always attempts to use the second derivative test and needs a backup for the cases where the test happens to be inconclusive. On the other hand, if one is going to construct a sign table for the second derivative to use Wu's result, would it not have been better to do the exact same thing on the first derivative and to use the first derivative test instead? After all, the whole point of the second derivative test is that we only need to rely on function evaluations at the critical points, without examining their surrounding neighborhoods. This leads to the following question: can we retain this local nature of the second derivative test and generalize it instead by considering the information provided by higher-order derivatives when evaluated only at the critical point?

In the present paper, we answer in the affirmative via Theorem 4.1 and Theorem 4.2. The underlying methodology is that when the second derivative test fails, we check higher-order derivatives, evaluated at the candidate point  $x_0$ , until a non-zero result is encountered. Then one of the above two theorems becomes applicable and the critical point can be classified. Interestingly, the proof of these theorems is based on an odd-order derivative generalization of the first derivative test which will be established by induction (see Lemma 3.2 and Lemma 3.3). Some readers may find these odd-order derivative generalizations as interesting as the main results. While the above theorems can be foreseen if one thinks in terms of a Taylor expansion, we prefer to provide direct proofs, circumventing the use of a Taylor expansion, which is typically taught much later down the road. The proofs are given in detail, without any logical leaps. We also illustrate, by example, the usefulness of quantifier notation towards writing complete and precise statements, proofs, and solutions to problems.

The paper is organized as follows. Preliminaries are given in section 2. The odd-order derivative generalization of the first derivative test is proved in section 3, and the generalization of the second derivative test is given in section 4. An example is given in section 5, and a brief discussion in section 6.

## 2. Preliminaries

We begin by reviewing preliminary results that need to be incorporated in the curriculum in order to be able to teach the proofs for the main results. Let  $f : A \rightarrow \mathbb{R}$  be a real-valued function with domain  $A \subseteq \mathbb{R}$  and let  $I \subseteq A$ . We use the notation  $f \uparrow I$  to state that  $f$  is strictly increasing on  $I$ , and  $f \downarrow I$  to state that  $f$  is strictly decreasing on  $I$ . The corresponding definitions are given by:

$$\begin{aligned} f \uparrow I &\iff \forall x_1, x_2 \in I : (x_1 < x_2 \implies f(x_1) < f(x_2)), \\ f \downarrow I &\iff \forall x_1, x_2 \in I : (x_1 < x_2 \implies f(x_1) > f(x_2)). \end{aligned}$$

As is well known, a strictly positive first derivative over a set implies that the function is strictly increasing over that set, and, similarly, a strictly negative first derivative over a set implies that the function is strictly decreasing over the same

set. For the proofs below we need a sharper version of this result which reads:

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) > 0 \\ f \text{ is continuous on } x = b \end{array} \right\} \implies f \uparrow (a, b),$$

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) > 0 \\ f \text{ is continuous on } x = a \end{array} \right\} \implies f \uparrow [a, b),$$

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) < 0 \\ f \text{ is continuous on } x = b \end{array} \right\} \implies f \downarrow (a, b),$$

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) < 0 \\ f \text{ is continuous on } x = a \end{array} \right\} \implies f \downarrow [a, b).$$

Here we use braces to represent *conjunction* between two or more statements (i.e. the logical “and”, as in  $p \wedge q$ ), with each statement given in a separate line. The idea is that, in general terms, differentiability and a non-zero derivative are required only for the interior of isolated intervals, whereas at the boundary points of the intervals, all that is required is continuity. The proof for these sharper theorems is the standard proof via the mean-value theorem which is already given in mainstream textbooks [4–8]. The key is to realize that the mean-value theorem itself requires differentiability only at the interior of the interval, with continuity being sufficient at the interval endpoints. Furthermore, after applying the mean-value theorem, the resulting first derivative will always be evaluated on the interior of the interval, therefore it is only necessary to know the sign of the first derivative at the open interval  $(a, b)$  for all cases. We note that the above results are needed to establish a strong version of the first derivative test (see Theorem 3.1), and furthermore in the proof of Theorem 4.2. By a similar line of reasoning, one can also claim the following additional results:

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) > 0 \\ f \text{ is continuous on } \{a, b\} \end{array} \right\} \implies f \uparrow [a, b],$$

$$\left. \begin{array}{l} \forall x \in (a, b) : f'(x) < 0 \\ f \text{ is continuous on } \{a, b\} \end{array} \right\} \implies f \downarrow [a, b].$$

However, these will not be needed for the proofs given below.

The proof of the second derivative test generalizations (i.e. Theorem 4.1 and Theorem 4.2) also require the following results from the theory of limits:

$$\lim_{x \rightarrow \sigma} f(x) > 0 \implies \exists \delta > 0 : \forall x \in N(\sigma, \delta) \cap A : f(x) > 0, \quad (1)$$

$$\lim_{x \rightarrow \sigma} f(x) < 0 \implies \exists \delta > 0 : \forall x \in N(\sigma, \delta) \cap A : f(x) < 0. \quad (2)$$

Here  $\sigma$  is a generalized accumulation point of the domain  $A = \text{dom}(f)$  of the function  $f$  (i.e.  $\sigma = x_0 \in \mathbb{R}$ , or  $\sigma = x_0^+$ , or  $\sigma = x_0^-$ , or  $\sigma = +\infty$ , or  $\sigma = -\infty$ ), and  $N(\sigma, \delta)$  is a generalized neighborhood of  $\sigma$  defined as:

$$N(\sigma, \delta) = \begin{cases} (x_0 - \delta, x_0 + \delta) - \{x_0\} & , \text{ if } \sigma = x_0 \\ (x_0 - \delta, x_0) & , \text{ if } \sigma = x_0^- \\ (x_0, x_0 + \delta) & , \text{ if } \sigma = x_0^+ \\ (1/\delta, +\infty) & , \text{ if } \sigma = +\infty \\ (-\infty, -1/\delta) & , \text{ if } \sigma = -\infty \end{cases} . \quad (3)$$

Note that for the case where  $\sigma$  represents a finite full or side limit,  $N(\sigma, \delta)$  is also known as a *deleted neighborhood*. It is worth noting that the converse of this result cannot be proven without weakening the inequalities into weak inequalities (i.e. greater-or-equal or smaller-or-equal). As was seen previously [9], the converse statements are needed to establish the theorems used to determine function convexity, but they do not become relevant to the problem at hand.

### 3. Generalizing the first derivative test

We begin with the precise statement of the first derivative test.

**Theorem 3.1:** *Let  $f : A \rightarrow \mathbb{R}$  be a function and let  $x_0 \in A$  be a point in the interior of the domain  $A$  of  $f$  such that it satisfies all of the following conditions:*

- (1)  $f$  differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$
- (2)  $f$  continuous on  $x_0$

Then, it follows that

$$\begin{aligned} \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0 \end{cases} &\implies x_0 \text{ is local minimum of } f, \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0 \end{cases} &\implies x_0 \text{ is local maximum of } f, \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0 \end{cases} &\implies x_0 \text{ is not extremum of } f, \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0 \end{cases} &\implies x_0 \text{ is not extremum of } f. \end{aligned}$$

Note that mainstream texts, like Stewart [4, 7, 8], are not careful in their formulation of the first derivative test with respect to clarifying the minimum differentiability and continuity requirements. On the other hand, the formulation of the first derivative test given by Apostol [10] is equivalent to the statement given above, and we may therefore rely on his derivation.

We now generalize the first derivative test in two steps. First, we show how the assumptions of the first derivative test can be derived from similar assumptions on the third derivative. In fact, the interesting result here is that, given all the needed differentiability and continuity assumptions, if the third derivative changes sign across a certain point in a certain way, then the first derivative changes sign across the same point in the same way. Second, we then generalize this by induction for all odd-order derivatives.

We begin with the following lemma linking the third derivative with the first derivative.

**Lemma 3.2:** *Let  $f : A \rightarrow \mathbb{R}$  be a function with domain  $A \subseteq \mathbb{R}$  and let  $x_0 \in A$  be a point interior to  $A$  such that it satisfies all of the following conditions:*

- (1)  $f$  is 3 times differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$
- (2)  $f$  is 2 times differentiable on  $x_0$
- (3)  $f''$  continuous on  $x_0$
- (4)  $f'(x_0) = f''(x_0) = 0$

Then, the statements below follow:

$$\begin{aligned} \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0 \end{cases} &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0, \end{cases} \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) < 0 \end{cases} &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0, \end{cases} \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0 \end{cases} &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0, \end{cases} \\ \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) < 0 \end{cases} &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) < 0. \end{cases} \end{aligned}$$

The proof of the lemma is based on the mean value theorem. Before giving the details, the careful reader may note that the function  $f$  itself does not appear anywhere in the statement of the lemma, and it will also not appear anywhere in the proof given below. Consequently, we could replace  $f'$  with  $g$ ,  $f''$  with  $g'$ , and  $f'''$  with  $g''$ , thereby shifting the degree of all multiple differentiations down by one differentiation, to obtain a somewhat stronger reformulation of the lemma. Although the function  $f$  does appear trivially on the differentiability conditions (1) and (2), it is easy to rewrite these conditions too in terms of a  $g \equiv f'$ . Furthermore, the very careful reader will realize that the reformulated lemma is in fact a special case of Yu's [3] generalized second derivative test, from which one can very quickly prove Yu's result in short order. On the other hand, we feel that the mathematical payoff from reformulating the lemma in this way does not justify the pedagogical penalty of the ensuing confusion, consequently we will retain and now prove the lemma as originally stated above:

**Proof:** It is sufficient to prove one of the above four statements, as the proof for the other three statements is very similar. So, with no loss of generality, let us assume that

$$\begin{cases} \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0. \end{cases}$$

Let  $x \in (x_0 - \delta, x_0)$  be given. From the given differentiability and continuity assumptions, we may apply the mean-value theorem to  $f'$  on the interval  $[x, x_0]$ . It follows that

$$\exists x_1 \in (x, x_0) : f''(x_1) = \frac{f'(x) - f'(x_0)}{x - x_0}. \quad (4)$$

Likewise, we apply the mean-value theorem to  $f''$  on the interval  $[x_1, x_0]$  and it follows that

$$\exists x_2 \in (x_1, x_0) : f'''(x_2) = \frac{f''(x_1) - f''(x_0)}{x_1 - x_0}. \quad (5)$$

Combining the above two equations we find that

$$\begin{aligned} f'(x) &= f'(x) - f'(x_0) && \text{[because } f'(x_0) = 0\text{]} \\ &= f''(x_1)(x - x_0) && \text{[via Eq. (4)]} \\ &= [f''(x_1) - f''(x_0)](x - x_0) && \text{[because } f''(x_0) = 0\text{]} \end{aligned}$$

$$\begin{aligned}
 &= f'''(x_2)(x_1 - x_0)(x - x_0) && \text{[via Eq. (5)]} \\
 &< 0,
 \end{aligned}$$

because  $f'''(x_2) < 0$  and  $x_1 - x_0 < 0$  and  $x - x_0 < 0$ . We conclude that  $\forall x \in (x_0 - \delta, x_0) : f'(x) < 0$ .

Now, let  $x \in (x_0, x_0 + \delta)$  be given. We apply the mean-value theorem on  $f'$  at the interval  $[x_0, x]$  and conclude that

$$\exists x_1 \in (x_0, x) : f''(x_1) = \frac{f'(x) - f'(x_0)}{x - x_0}.$$

We also apply the mean-value theorem on  $f''$  at the interval  $[x_0, x_1]$  and find that

$$\exists x_2 \in (x_0, x_1) : f'''(x_2) = \frac{f''(x_1) - f''(x_0)}{x_1 - x_0}.$$

From the above equations, it follows again that

$$f'(x) = f'''(x_2)(x_1 - x_0)(x - x_0) > 0,$$

because now we have  $f'''(x_2) > 0$  and  $x_1 - x_0 > 0$  and  $x - x_0 > 0$ , instead. We conclude that  $\forall x \in (x_0, x_0 + \delta) : f'(x) > 0$ .

Similar proofs can be given for the other three statements.  $\square$

Combining the previous lemma with the first derivative test we may now prove the following generalization of the first derivative test by induction. Here we will use the definition  $[n] = \{1, 2, 3, \dots, n\}$  to represent the set of all integers from 1 to  $n$ .

**Lemma 3.3:** *Let  $f : A \rightarrow \mathbb{R}$  be a function, let  $x_0 \in A$  be an interior point, and let  $n \in \mathbb{N} - \{0\}$  such that all of the following conditions are satisfied:*

- (1)  $f$  is  $2n + 1$  times differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$
- (2)  $f$  is  $2n$  times differentiable on  $x_0$
- (3)  $f^{(2n)}$  continuous on  $x_0$

Then it follows that:

$$\left. \begin{aligned} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) > 0 \end{aligned} \right\} \implies x_0 \text{ is local minimum of } f,$$

$$\left. \begin{aligned} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) < 0 \end{aligned} \right\} \implies x_0 \text{ is local maximum of } f,$$

$$\left. \begin{aligned} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) > 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) > 0 \end{aligned} \right\} \implies x_0 \text{ is not extremum of } f,$$

$$\left. \begin{aligned} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) < 0 \end{aligned} \right\} \implies x_0 \text{ is not extremum of } f.$$

**Proof:** We begin with a proof of the first statement. For the case  $n = 1$ , we note

that

$$\left\{ \begin{array}{l} f'(x_0) = f''(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f'''(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'''(x) > 0 \end{array} \right\} \implies \left\{ \begin{array}{l} f'(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f'(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f'(x) > 0 \end{array} \right\} \\ \implies x_0 \text{ is local minimum of } f.$$

Here, the first implication step uses Lemma 3.2 and the second implication step uses the first derivative test. This sequence of implications proves the theorem for  $n = 1$ .

For  $n = k$ , let us assume that it has been shown that

$$\left. \begin{array}{l} \forall \alpha \in [2k] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2k+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2k+1)}(x) > 0 \end{array} \right\} \implies x_0 \text{ is local minimum of } f.$$

Then, for the case  $n = k + 1$ , we have

$$\left\{ \begin{array}{l} \forall \alpha \in [2k + 2] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2k+3)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2k+3)}(x) > 0 \end{array} \right\} \implies \left\{ \begin{array}{l} \forall \alpha \in [2k] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2k+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2k+1)}(x) > 0 \end{array} \right\} \\ \implies x_0 \text{ is local minimum of } f.$$

Here, the first step applies Lemma 3.2 to  $f^{(2k)}$ , and the second step uses the inductive hypothesis. Therefore, the statement

$$\left. \begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ \forall x \in (x_0 - \delta, x_0) : f^{(2n+1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n+1)}(x) > 0 \end{array} \right\} \implies x_0 \text{ is local minimum of } f,$$

follows by induction for all  $n \in \mathbb{N} - \{0\}$ . A similar argument proves the remaining statements.  $\square$

#### 4. Generalization of the 2nd derivative test

Using Lemma 3.3, we will now derive two theorems that constitute the proposed generalization of the second derivative test. The overall idea is that if it should happen that at a critical point  $x_0 \in A$ , in the domain  $A$  of some function  $f$ , we have  $f'(x_0) = f''(x_0) = 0$ , we continue checking higher-order derivatives such as  $f'''(x_0), \dots, f^{(n)}(x_0)$  until we encounter a non-zero derivative. It should be easy to foresee, via Taylor expansion, that unless a function is locally constant around  $x_0$ , sooner or later we should find some  $n \in \mathbb{N} - \{0\}$  such that  $f^{(n)}(x_0) \neq 0$ , with all lower-order derivatives being equal to zero. If  $n$  is even, then Theorem 4.1 will classify the point  $x_0$  as a local minimum or maximum, depending on the sign of  $f^{(n)}(x_0)$ . If  $n$  is odd, then Theorem 4.2 shows that  $x_0$  is neither a local minimum nor a local maximum. The proof of Theorem 4.1 closely follows the same argument one uses to prove the second derivative test, except that the argument is interfaced with Lemma 3.3 instead of the first derivative test. The proof of Theorem 4.2 requires a trickier argument that has to take advantage of the generalized monotonicity

criteria, which are probably not emphasized in usual coursework, as they are not needed in the proof of the original second derivative test.

Let us now state and prove these theorems below:

**Theorem 4.1:** *Let  $f : A \rightarrow \mathbb{R}$  be a function, let  $x_0 \in A$  be an interior point, and let  $n \in \mathbb{N} - \{0\}$ . We assume that all of the following statements are satisfied:*

- (1)  $f$  is  $2n - 1$  times differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$
- (2)  $f$  is  $2n$  times differentiable on  $x_0$

Then, it follows that

$$\left. \begin{array}{l} \forall \alpha \in [2n - 1] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n)}(x_0) > 0 \end{array} \right\} \implies x_0 \text{ is local minimum of } f,$$

$$\left. \begin{array}{l} \forall \alpha \in [2n - 1] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n)}(x_0) < 0 \end{array} \right\} \implies x_0 \text{ is local maximum of } f.$$

**Proof:** We begin with establishing the first statement. Assume that  $\forall \alpha \in [2n - 1] : f^{(\alpha)}(x_0) = 0$  and  $f^{(2n)}(x_0) > 0$ . First, we note that

$$\begin{aligned} f^{(2n)}(x_0) &= \lim_{x \rightarrow x_0} \frac{f^{(2n-1)}(x) - f^{(2n-1)}(x_0)}{x - x_0} && \text{[by definition]} \\ &= \lim_{x \rightarrow x_0} \frac{f^{(2n-1)}(x)}{x - x_0} > 0 && \text{[via } f^{(2n-1)}(x_0) = 0] \\ &\implies \exists \delta > 0 : \forall x \in N(x_0, \delta) : \frac{f^{(2n-1)}(x)}{x - x_0} > 0 \\ &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n-1)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n-1)}(x) > 0. \end{cases} \end{aligned}$$

The first implication uses the limit property given by Eq. (1), where the differentiability assumptions given by the theorem allows us to replace  $N(x_0, \delta) \cap A$  with  $N(x_0, \delta)$ . The second implication follows from noting that  $x - x_0$  and  $f^{(2n-1)}(x_0)$  have the same sign, and also that  $x - x_0$  is negative when  $x \in (x_0 - \delta, x_0)$  and positive when  $x \in (x_0, x_0 + \delta)$ . Combining the above result with the assumption  $\forall \alpha \in [2n - 1] : f^{(\alpha)}(x_0) = 0$  via the Lemma 3.3, we conclude that  $x_0$  is a local minimum. It is worth noting that Lemma 3.3 only requires, in this case, that  $f$  be  $2n - 1$  times differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$ ,  $f$  be  $2n - 2$  times differentiable on  $x_0$ , and  $f^{(2n-2)}$  be continuous on  $x_0$ . These conditions are all covered by the assumptions used by this theorem.

The next statement can be derived via a similar argument, noting that now  $f^{(2n-1)}(x_0)$  and  $x - x_0$  will have opposite signs. As a result, via a similar argument, it will follow at the end that  $x_0$  is a local maximum.  $\square$

**Theorem 4.2:** *Let  $f : A \rightarrow \mathbb{R}$  be a function, let  $x_0 \in A$  be an interior point, and let  $n \in \mathbb{N} - \{0\}$ . We assume that all of the following statements are satisfied:*

- (1)  $f$  is  $2n$  times differentiable on  $(x_0 - \delta, x_0) \cup (x_0, x_0 + \delta)$
- (2)  $f$  is  $2n + 1$  times differentiable on  $x_0$

Then it follows that

$$\left. \begin{array}{l} \forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0 \\ f^{(2n+1)}(x_0) \neq 0 \end{array} \right\} \implies x_0 \text{ is not extremum of } f.$$



**Proof:** Let us assume with no loss of generality that  $f^{(2n+1)}(x_0) > 0$ . Then it follows that

$$\begin{aligned}
 f^{(2n+1)}(x_0) &= \lim_{x \rightarrow x_0} \frac{f^{(2n)}(x) - f^{(2n)}(x_0)}{x - x_0} && \text{[by definition]} \\
 &= \lim_{x \rightarrow x_0} \frac{f^{(2n)}(x)}{x - x_0} > 0 && \text{[via } f^{(2n)}(x_0) = 0\text{]} \\
 &\implies \exists \delta > 0 : \forall x \in N(x_0, \delta) : \frac{f^{(2n)}(x)}{x - x_0} > 0 \\
 &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n)}(x) < 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n)}(x) > 0 \end{cases} \\
 &\implies \begin{cases} f^{(2n-1)} \downarrow (x_0 - \delta, x_0) \\ f^{(2n-1)} \uparrow [x_0, x_0 + \delta) \end{cases} \\
 &\implies \begin{cases} \forall x \in (x_0 - \delta, x_0) : f^{(2n-1)}(x) > f^{(2n-1)}(x_0) = 0 \\ \forall x \in (x_0, x_0 + \delta) : f^{(2n-1)}(x) > f^{(2n-1)}(x_0) = 0. \end{cases}
 \end{aligned}$$

The first two implications follow the same line of reasoning as in the preceding proof. Specifically, the first implication uses the limit properly given by Eq. (1), and the second implication is based on noting that  $f^{(2n)}(x_0)$  and  $x - x_0$  have the same sign. The third implication uses the generalized monotonicity theorems given in section 2, and the fourth implication uses the definition of monotonicity. Combining the above result with the assumption  $\forall \alpha \in [2n] : f^{(\alpha)}(x_0) = 0$ , we conclude via Lemma 3.3 that  $x_0$  is neither a local minimum nor a local maximum. It is important to note that in the third implication step we go from open to closed intervals, and doing so can only be justified via the generalized monotonicity theorems. This transition to closed intervals, on the  $x_0$  side, is in turn necessary to complete the fourth implication.  $\square$

## 5. A trigonometric example

Polynomial functions do not make for interesting application problems for the second derivative test and the generalizations given by Theorem 4.1 and Theorem 4.2. If one has found all of the critical points that are candidates for local minimum and local maximum, it follows that a factorization of the first derivative into linear and quadratic factors has also been found, and it is very easy to apply the first derivative test using the given factorization. Trigonometric functions, on the other hand, lead to an infinite set of critical points, and using the first derivative test on such problems tends to be cumbersome.

Simple trigonometric functions can be usually handled via the second derivative test with relative efficiency, but it is possible to define trigonometric functions where the second derivative test fails. In this section we will consider such an example. A unique challenge posed by trigonometric problems is that some set theory is needed to organize and keep track of the critical points. Consequently, the example given below could be too complicated for a regular calculus course, but may be interesting for an honors calculus course or for a gateway course intended to introduce mathematical proofs and rigorous thinking. Also note that the derivative calculations are quite cumbersome, so students can be encouraged to use computer algebra software for problems at this level of complexity (the author used Maxima on an Android phone).

**Example 5.1** Find all local minima and maxima of the function defined by  $f(x) = \sin(4x) \cos^4(x)$ .

**Solution :** The first derivative of  $f(x)$  is given by

$$f'(x) = 4 \cos^3 x [\cos(4x) \cos x - \sin(4x) \sin x] = 4 \cos^3 x \cos(5x).$$

Since the domain of the function is  $\text{dom}(f) = \mathbb{R}$ , we can find all candidates for local minimum and local maximum by solving

$$\begin{aligned} f'(x) = 0 &\iff 4 \cos^3 x \cos(5x) = 0 \iff \cos x = 0 \vee \cos(5x) = 0 \\ &\iff \left( \exists k \in \mathbb{Z} : x = \frac{\pi}{2} + k\pi \right) \vee \left( \exists k \in \mathbb{Z} : 5x = \frac{\pi}{2} + k\pi \right) \\ &\iff \exists k \in \mathbb{Z} : \left( x = \frac{\pi}{2} + k\pi \vee x = \frac{\pi}{10} + \frac{k\pi}{5} \right). \end{aligned}$$

Let us now group the solutions into two sets given by

$$\begin{aligned} S_1 &= \{x \in \mathbb{R} \mid \cos x = 0\} = \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\}, \\ S_2 &= \{x \in \mathbb{R} \mid \cos(5x) = 0\} = \left\{ \frac{\pi}{10} + \frac{k\pi}{5} \mid k \in \mathbb{Z} \right\}. \end{aligned}$$

The second derivative is given by

$$f''(x) = \dots = -4 \cos^2 x [5 \cos x \sin(5x) + 3 \sin x \cos(5x)],$$

and it is easy to see that for all candidate points in  $S_1$ , the second derivative test fails. Specifically, by definition we know that  $\forall x \in S_1 : \cos x = 0$  and consequently it follows that  $\forall x \in S_1 : f''(x) = 0$ . Another complication is that the sets  $S_1$  and  $S_2$  seem to have elements in common. It is therefore necessary to reorganize our sets of possible critical points, so we can then consider them on a case by case basis.

We claim that  $S_1 \subseteq S_2$ . We prove the claim as follows:

$$\begin{aligned} x \in S_1 &\implies \exists k \in \mathbb{Z} : x = \frac{\pi}{2} + k\pi = \frac{\pi}{10} + \frac{4\pi}{10} + k\pi = \frac{\pi}{10} + \frac{2\pi}{5} + k\pi \\ &= \frac{\pi}{10} + \frac{(5k+2)\pi}{5} \\ &\implies \exists \lambda \in \mathbb{Z} : x = \frac{\pi}{10} + \frac{\lambda\pi}{5} \implies x \in S_2. \end{aligned}$$

This concludes the proof of the claim.

Now, let us separate the set  $S_2$  into 5 disjoint subsets  $S_{2,\alpha}$  defined as:

$$\forall \alpha \in [5] : S_{2,\alpha} = \left\{ \frac{\pi}{10} + \frac{(5k+\alpha)\pi}{5} \mid k \in \mathbb{Z} \right\}.$$

For  $\alpha = 2$ , we have already shown that  $S_1 \subseteq S_{2,2}$ . We will now claim that  $S_1 = S_{2,2}$  and  $S_1 \cap S_{2,\alpha} = \emptyset$ ,  $\forall \alpha \in \{1, 3, 4, 5\}$ . To prove the claim, we first note that

$$x \in S_{2,\alpha} \implies \exists k \in \mathbb{Z} : x = \frac{\pi}{10} + \frac{(5k+\alpha)\pi}{5} = \frac{\pi}{10} + \frac{(5k+2)\pi}{5} + \frac{(\alpha-2)\pi}{5}$$

$$= \frac{\pi}{2} + k\pi + \frac{(\alpha - 2)\pi}{5} = \frac{\pi}{2} + \left[ k + \frac{\alpha - 2}{5} \right] \pi.$$

We may now distinguish between the following two cases:

- **Case 1:** If  $\alpha = 2$ , then it follows that

$$x = \frac{\pi}{2} + k\pi \implies x \in S_1.$$

Consequently, we have established that  $S_{2,2} \subseteq S_1$ , and therefore  $S_1 = S_{2,2}$ .

- **Case 2:** For  $\alpha \in \{1, 3, 4, 5\}$ , we note that

$$\begin{aligned} \left( \forall k \in \mathbb{Z} : k + \frac{\alpha - 2}{5} \notin \mathbb{Z} \right) &\implies \left( \forall \lambda \in \mathbb{Z} : x \neq \frac{\pi}{2} + \lambda\pi \right) \\ &\implies x \notin S_1. \end{aligned}$$

It follows that  $S_1 \cap S_{2,\alpha} = \emptyset$ ,  $\forall \alpha \in \{1, 3, 4, 5\}$ .

This concludes the proof of the claim.

From the above results, we see that all critical points are in the sets  $S_{2,\alpha}$  for  $\alpha \in [5]$ , and that all critical points for which the second derivative test fails, are located in  $S_{2,2}$ . We will therefore first classify the critical points for  $\alpha \in \{1, 3, 4, 5\}$  using the second derivative test, and consider the case  $\alpha = 2$  separately. First, we note that in general,

$$x \in S_{2,\alpha} \implies \exists k \in \mathbb{Z} : x = \frac{\pi}{10} + \frac{(5k + \alpha)\pi}{5},$$

and it follows that  $\cos(5x) = 0$  and

$$\begin{aligned} \sin(5x) &= \sin \left[ 5 \left( \frac{\pi}{10} + \frac{(5k + \alpha)\pi}{5} \right) \right] = \sin \left( \frac{\pi}{2} + (5k + \alpha)\pi \right) \\ &= \cos((5k + \alpha)\pi) = (-1)^{5k + \alpha}, \end{aligned}$$

and,

$$\begin{aligned} \cos x &= \cos \left( \frac{\pi}{10} + \frac{(5k + \alpha)\pi}{5} \right) = \cos \left( \frac{\pi}{2} + k\pi + \frac{(\alpha - 2)\pi}{5} \right) \\ &= (-1)^k \cos \left( \frac{\pi}{2} + \frac{(\alpha - 2)\pi}{5} \right) = (-1)^{k+1} \sin \left( \frac{(\alpha - 2)\pi}{5} \right), \end{aligned}$$

and using the above equations, the second derivative evaluates to

$$\begin{aligned} f''(x) &= -4 \cos^2 x [5 \cos x \sin(5x) + 3 \sin x \times 0] = -20 \cos^3 x \sin(5x) \\ &= -20 \left[ (-1)^{k+1} \sin \left( \frac{(\alpha - 2)\pi}{5} \right) \right]^3 \times (-1)^{5k + \alpha} \\ &= -20 (-1)^{5k + \alpha} (-1)^{3k + 3} \sin^3 \left( \frac{(\alpha - 2)\pi}{5} \right) \\ &= -20 (-1)^{8k + \alpha + 3} \sin^3 \left( \frac{(\alpha - 2)\pi}{5} \right) = 20 (-1)^{\alpha} \sin^3 \left( \frac{(\alpha - 2)\pi}{5} \right) \end{aligned}$$

On a case-by-case basis, it follows that

$$\begin{aligned} \text{For } a = 1 : \quad & f''(x) = 20(-1)^1 \sin^3(-\pi/5) > 0 \\ & \implies \forall x \in S_{2,1} : x \text{ local minimum;} \\ \text{For } a = 3 : \quad & f''(x) = 20(-1)^3 \sin^3(\pi/5) < 0 \\ & \implies \forall x \in S_{2,3} : x \text{ local maximum;} \\ \text{For } a = 4 : \quad & f''(x) = 20(-1)^4 \sin^3(2\pi/5) > 0 \\ & \implies \forall x \in S_{2,4} : x \text{ local minimum;} \\ \text{For } a = 5 : \quad & f''(x) = 20(-1)^5 \sin^3(3\pi/5) < 0 \\ & \implies \forall x \in S_{2,5} : x \text{ local maximum.} \end{aligned}$$

On the other hand, for the case  $\alpha = 2$ , we find

$$f''(x) = 20(-1)^2 \sin^3(0) = 0,$$

so the second derivative test fails. The next two derivatives are given by

$$\begin{aligned} f'''(x) &= 8 \cos x [15 \cos x \sin x \sin(5x) + 3 \sin^2 x \cos(5x) - 14 \cos^2 x \cos(5x)], \\ f^{(4)}(x) &= -8[(45 \cos x \sin^2 x - 85 \cos^3 x) \sin(5x) + (3 \sin^3 x - 123 \cos^2 \sin x) \cos(5x)]. \end{aligned}$$

However, since for all  $x \in S_{2,2}$  we have  $\cos x = 0$  and  $\cos(5x) = 0$ , both derivatives simplify to  $f'''(x) = 0$  and  $f^{(4)}(x) = 0$ . So, we move on to the fifth derivative, which reads and simplifies as:

$$\begin{aligned} f^{(5)}(x) &= 32[(15 \sin^3 x - 240 \cos^2 x \sin x) \sin(5x) \\ &\quad + (137 \cos^3 x - 120 \cos x \sin^2 x) \cos(5x)] \\ &= (32 \times 15) \sin^3 x \sin(5x) = 480 \sin^3 x \sin(5x). \end{aligned}$$

Here we have already used  $\cos x = 0$ , and  $\cos(5x) = 0$ . For the remainder of the calculation we need  $\sin(5x) = (-1)^{5k+2} = (-1)^{5k}$ , which was shown previously, and we also need  $\sin x$ , which is calculated as follows:

$$\sin x = \sin(k\pi + \pi/2) = (-1)^k \sin(\pi/2) = (-1)^k.$$

Consequently, the fifth derivative is given by

$$f^{(5)}(x) = 480[(-1)^k]^3 (-1)^{5k} = 480(-1)^{3k+5k} = 480(-1)^{8k} = 480 \neq 0.$$

Using Theorem 4.2, we conclude that for  $x \in S_{2,2}$

$$\begin{cases} \forall \alpha \in [4] : f^{(\alpha)}(x) = 0 \\ f^{(5)}(x) \neq 0 \end{cases} \implies x \text{ not a local extremum.}$$

To summarize, we see that the elements of  $S_{2,1} \cup S_{2,4}$  are local minima, the elements of  $S_{2,3} \cup S_{2,5}$  are local maxima, and the elements of  $S_{2,2}$  are neither local minima nor local maxima. □

## 6. Conclusion

The main advantage of the generalization of the second derivative test given by Theorem 4.1 and Theorem 4.2 is that they are local results where all relevant derivatives need only be evaluated at the critical points. The theorems are robust, since any non-constant analytic function will eventually yield a non-zero high-order derivative at each critical point. If it were the case that the entire countable set of higher-order derivatives evaluate to zero, for some point, then, via a Taylor expansion, it follows that the function would have to be locally constant in a neighborhood of that point. We have also seen that, combined with computer algebra software, these theorems provide a very powerful methodology for locating and classifying the local minima and local maxima of trigonometric functions. The generalized theorems could be mentioned in an honors calculus course and used to formulate projects that leverage computer algebra systems. They can also be made relevant in a proof class or even a real analysis class. To that end, the construction of a decent exercise set leveraging these theorems would be worth exploring in future articles.

## References

- [1] M. Nerenberg, *The second derivative test for constrained extremum problems*, Int. J. Math. Ed. Sci. Tech. 22 (1991), pp. 303–308.
- [2] R. Vyborny and R. Nester, *First derivative tests for extrema*, Int. J. Math. Ed. Sci. Tech. 20 (1989), pp. 755–762.
- [3] Y. Wu, *Improved second derivative test for relative extrema*, Int. J. Math. Ed. Sci. Tech. 38 (2007), pp. 1121–1123.
- [4] J. Stewart *Single Variable Calculus: Early Transcendentals*, Brooks/Cole Publishing Company, Pacific Grove CA, 1999.
- [5] D. Varberg, E. Purcell, and S. Rigdon *Calculus*, Pearson Prentice Hall, New Jersey NJ, 2007.
- [6] J. Haas, M. Weir, and G. Thomas *University Calculus. Early Transcendentals*, Pearson Education, Boston MA, 2012.
- [7] J. Stewart *Essential Calculus. Early Transcendentals*, Thomson Brooks/Cole, Belmont CA, 2007.
- [8] J. Stewart *Calculus*, Thomson Learning, Belmont CA, 2003.
- [9] E. Gkioulekas, *On equivalent characterizations of convexity of functions*, Int. J. Math. Ed. Sci. Tech. (2012) DOI:10.1080/0020739X.2012.703336.
- [10] T. Apostol *Calculus. I*, John Wiley & Sons, New York NY, 1967.