Programming Research Group

INFORMATIC VS. CLASSICAL DIFFERENTIABILITY ON THE REAL LINE

Keye Martin and Joël Ouaknine

PRG-RR-03-04

Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD

Abstract

We study the relationship between informatic and classical differentiation on the real line. The former arises when considering the interval domain over the reals equipped with the Lebesgue measurement. We show that informatic differentiation is a strict generalization of its classical counterpart, and wonder if it can provide a springboard towards extending techniques and results from calculus to certain non classically differentiable functions.

1 Introduction

Informatic differentiation arises in the study of domains and measurements [3]. Given a selfmap on a domain, its informatic derivative measures the rate at which the content of the output changes with respect to the content of the input. This, for instance, can be used to record how quickly a selfmap converges to a fixed point, and sheds light on whether a fixed point is a local attractor or not. These ideas have been used to study the complexity of algorithms [5] and in the search for fixed points and zeros of functions [4].

When specialized to the interval domain over the reals and continuous real-valued functions, informatic differentiation bears strong similarities with its classical counterpart. In this paper, we make precise and examine some of these similarities, and show that informatic differentiation is a strict generalization of classical differentiation. One of the most promising consequences of this observation is the possibility to extend certain results from calculus to functions that do not necessarily admit classical derivatives.

2 Background

We shall be working exclusively with intervals of the real line, so knowledge of domain theory is not required to read this paper. Still, let us hint at the general idea of informatic differentiation (complete details are in [3]): Given a domain and measurement (D, μ), the informatic derivative of $f : D \to D$ at p with respect to μ is

$$df_\mu(p) \mathrel{\widehat{=}} \lim_{x \to p} \frac{\mu f(x) - \mu f(p)}{\mu x - \mu p}$$

where the limit, if it exists, is taken in the μ topology on D. In the case we are interested in, D will be the *interval domain* $\mathbf{I}\mathbb{R} \mathrel{\widehat{=}} \{[a, b] \mid a, b \in \mathbb{R} \wedge a \leqslant b\}$ of compact intervals over the reals, and μ will be the standard measure of length, $\mu[a, b] = b - a$. For $p \in \mathbb{R}$ we will denote the interval $[p, p]$ by $[p]$. Thus, $\mu[p] = 0$.

Now, any continuous map $f : \mathbb{R} \to \mathbb{R}$ has a canonical extension $\bar{f} : \mathbf{I}\mathbb{R} \to \mathbf{I}\mathbb{R}$. The informatic derivative of $\bar{f}$ at $[p]$ with respect to μ is then

$$d\bar{f}_\mu[p] \mathrel{\widehat{=}} \lim_{I \to [p]} \frac{\mu \bar{f}(I)}{\mu I}$$

where the limit is taken over all non-trivial intervals $I \supset [p]$ with $\mu I \to 0$.

The informatic and classical derivatives of a function are closely related as the following theorem from [3] shows:

Theorem 2.1 *Let $f \in C(\mathbb{R})$ and let $a, b \in \mathbb{R}$ with $a < b$. Then for $p \in \mathbb{R}$,*

(i) *$d\bar{f}_\mu[p] \geqslant 0$ whenever it exists.*

(ii) *If $f'(p)$ exists, then $d\bar{f}_\mu[p] = |f'(p)|$.*

(iii) *$d\bar{f}_\mu[p] = 0$ iff $f'(p) = 0$.*

(iv) *If $d\bar{f}_\mu[x] > 0$ for all $x \in (a, b)$, then f' exists on (a, b).*

(v) *$f \in C^1(\mathbb{R})$ iff $d\bar{f}_\mu \in C(\mathbb{R})$.*

Proof. (Sketch.) (i), (ii), and (iii) are straightforward. For (iv), one first shows that, whenever f achieves a local extremum at some point $x \in \mathbb{R}$, then the only value that $d\bar{f}_\mu[x]$ can take, if it exists, is zero (cf. Example 2.2). Consequently, if $d\bar{f}_\mu[x] > 0$ for all $x \in (a, b)$, then f is monotone over (a, b), from which the result easily follows. Lastly, for (v), note that if $d\bar{f}_\mu$ is continuous at p, then $f'(p)$ must exist: either $df_\mu(p) = 0$, in which case $f'(p) = 0$, or $df_\mu(p) > 0$, and then by continuity $df_\mu > 0$ on an open interval about p. $\square$

Example 2.2 The functions $f(x) = x$ and $g(x) = -x$ both have informatic derivative 1 at $p = 0$. On the other hand, the function $h(x) = |x|$ has no informatic derivative at $p = 0$: indeed, when I is of the form $[0, x]$, the ratio $\mu\bar{h}(I)/\mu I = 1$, whereas when I is of the form $[-x, x]$, the ratio $\mu\bar{h}(I)/\mu I = 1/2$, which shows that the limit as $I \to [0]$ cannot exist.

3 Defining the Sign of the Informatic Derivative

It may seem strange that $d\bar{f}_\mu$ can capture differentiation without explicit reference to the *sign* of the derivative. The sign of the derivative, after all, is an important piece of qualitative information. We now show that the existence of an informatic derivative implicitly contains a description of the sign—even if f is not classically differentiable.

Lemma 3.1 *Let $f \in C(\mathbb{R})$ and $p \in \mathbb{R}$ be such that $d\bar{f}_\mu[p]$ exists. Then either*

$$\text{(i)} \lim_{e \to p^+} \frac{\max \bar{f}([p, e]) - f(p)}{e - p} = 0 \qquad \textit{or} \qquad \text{(ii)} \lim_{e \to p^+} \frac{\min \bar{f}([p, e]) - f(p)}{e - p} = 0 \ .$$

Moreover, $d\bar{f}_\mu[p] = 0$ iff both (i) *and* (ii) *hold, and $d\bar{f}_\mu[p] > 0$ iff exactly one of* (i) *and* (ii) *holds.*

Proof. The only statement that is not immediate is that at least one of (i) and (ii) must always hold. Without loss of generality, it suffices to consider the case $p = 0$, $f(0) = 0$, and $d\bar{f}_\mu[0] > 0$.

For $e > 0$, let $M^+(e) = (\max \bar{f}([0, e]))/e$ and $m^+(e) = (\min \bar{f}([0, e]))/e$. Suppose that neither (i) nor (ii) holds. Then there exists $q > 0$ and two decreasing sequences

$\langle e_1, e_2, \ldots \rangle$ and $\langle e'_1, e'_2, \ldots \rangle$ of positive reals, both tending to zero, such that, for all i, $M^+(e_i) > q$ and $m^+(e'_i) < -q$. Note that this entails that $d\bar{f}_\mu[0] \geqslant q$.

Let $\varepsilon < q/12$ be a positive real number. Find δ such that, for any non-trivial interval $I \subseteq \mathbb{R}$ containing 0 with $\mu(I) < \delta$, $|\mu\bar{f}(I)/\mu I - d\bar{f}_\mu[0]| < \varepsilon$.

We claim that there exists some $0 < e < \delta/2$ such that both $M^+(e) > q/3$ and $m^+(e) < -q/3$. To see this, choose i large enough so that $e_i, e'_i < \delta/2$. If neither e_i nor e'_i satisfies the condition, the continuous function $g(x) = \max \bar{f}([0,x]) + \min \bar{f}([0,x])$ must change signs between e_i and e'_i. The intermediate value theorem therefore yields a value e between e_i and e'_i at which g vanishes: $M^+(e) = -m^+(e) = (1/2)(\mu\bar{f}([0,e])/\mu[0,e])$. Note that $\mu\bar{f}([0,e])/\mu[0,e] > d\bar{f}_\mu[0] - \varepsilon > 11q/12$, and thus $M^+(e) > 11q/24 > q/3$. Likewise, $m^+(e) < -q/3$.

Let $M^-(e) = (\max \bar{f}([-e,0]))/e$ and $m^-(e) = (\min \bar{f}([-e,0]))/e$. Notice that we cannot both have $M^-(e) \geqslant M^+(e)$ and $m^-(e) \leqslant m^+(e)$, otherwise we would have $\mu\bar{f}([-e,e]) = \mu\bar{f}([-e,0])$ which would quickly lead to a contradiction on account of the inequalities $|\mu\bar{f}([-e,e])/\mu[-e,e] - d\bar{f}_\mu[0]| < \varepsilon$ and $|\mu\bar{f}([-e,0])/\mu[-e,0] - d\bar{f}_\mu[0]| < \varepsilon$. Likewise, we cannot both have $M^+(e) \geqslant M^-(e)$ and $m^+(e) \leqslant m^-(e)$. Without loss of generality let us therefore assume that $m^-(e) \leqslant m^+(e) < M^-(e) \leqslant M^+(e)$.

We now have

$$
\begin{aligned}
\frac{\mu\bar{f}([-e,e])}{\mu[-e,e]} = \frac{M^+(e) - m^-(e)}{2} = \\
\frac{(M^+(e) - m^+(e)) + (M^-(e) - m^-(e)) - (M^-(e) - m^+(e))}{2} < \\
\frac{(d\bar{f}_\mu[0] + \varepsilon) + (d\bar{f}_\mu[0] + \varepsilon) - (0 + q/3)}{2} = \\
d\bar{f}_\mu[0] + \varepsilon - q/6 < d\bar{f}_\mu[0] + q/12 - q/6 = d\bar{f}_\mu[0] - q/12 < d\bar{f}_\mu[0] - \varepsilon \ .
\end{aligned}
$$

In short, we get $\mu\bar{f}([-e,e])/\mu[-e,e] < d\bar{f}_\mu[0] - \varepsilon$, contradicting the hypothesis $|\mu\bar{f}([-e,e])/\mu[-e,e] - d\bar{f}_\mu[0]| < \varepsilon$. $\square$

There is then an underlying trichotomy at work: either we have exactly one of (i) and (ii), or $d\bar{f}_\mu[p] = 0$. This implicitly defines a quantity called the *sign* of $d\bar{f}_\mu$ at $[p]$, denoted $\mathrm{sgn}(d\bar{f}_\mu)(p)$, formally given by

$$
\mathrm{sgn}(d\bar{f}_\mu)(p) \mathrel{\widehat{=}} \begin{cases} +1 & \text{if only (ii)} \\ -1 & \text{if only (i)} \\ 0 & \text{otherwise} \end{cases} .
$$

The informatic derivative on the real line is a *signed quantity* in disguise.

Definition 3.2 For $f \in C(\mathbb{R})$ and $p \in \mathbb{R}$, we define

$$
df_\mu(p) \mathrel{\widehat{=}} \mathrm{sgn}(d\bar{f}_\mu)(p) \cdot d\bar{f}_\mu[p]
$$

whenever $d\bar{f}_\mu[p]$ exists.

Theorem 3.3 *Let $f \in C(\mathbb{R})$ and $p \in \mathbb{R}$ be such that $df_\mu(p)$ exists.*

(i) *$df_\mu(p) = f'(p)$ whenever $f'(p)$ exists.*

(ii) *If $df_\mu(p) > 0$, then for all sufficiently small non-trivial intervals $I = [p-\delta, p+\delta]$, if f achieves its minimum over I at x^- and its maximum over I at x^+, then $x^- < p < x^+$.*

(iii) *If $df_\mu(p) < 0$, then for all sufficiently small non-trivial intervals $I = [p-\delta, p+\delta]$, if f achieves its minimum over I at x^- and its maximum over I at x^+, then $x^+ < p < x^-$.*

Proof. (Sketch.) (i) is immediate. For (ii) and (iii), first apply Lemma 3.1 to $g(x) = f(2p - x)$, and then show that $\mathrm{sgn}(d\bar{g}_\mu)(p) = -\mathrm{sgn}(d\bar{f}_\mu)(p)$. □

A classical derivative can thus be factored into two more primitive ideas: (1) a sign, which gives qualitative information about the shape of the curve f, and (2) a number $d\bar{f}$, which expresses magnitude of rate of change.

4 Informatic vs. Classical Differentiability

Despite the close connection between the two notions, we now show that informatic differentiation is strictly more general than classical differentiation.

Let $l : [0,1] \to \mathbb{R}$ be $l(x) = \ln(x+1)$. Note that $0 < l(x) < x$ for all $x \in (0,1]$. Define a sequence $\langle a_1, a_2, \ldots \rangle$ as follows: $a_1 = 1$, and $a_{i+1} = l(a_i)$.

Lemma 4.1 *The sequence $\langle a_1, a_2, \ldots \rangle$ is strictly decreasing and converges to zero.*

Proof. We have $l(1) < 1$, i.e., $a_2 < a_1$. Assuming that $a_{n+1} < a_n$, applying l (which is strictly increasing) to both sides yields $a_{n+2} = l(a_{n+1}) < l(a_n) = a_{n+1}$. By induction, the sequence is hence strictly decreasing. Moreover, it is bounded below by 0, and thus has a limit, which by continuity must remain unchanged under l; 0 is the only such point. □

Next, define a sequence $\langle b_1, b_2, \ldots \rangle$ by picking b_i such that $a_{i+1} < b_i < a_i$. It is clear that the sequences $\langle b_1, b_2, \ldots \rangle$ and $\langle a_1, b_1, a_2, b_2, \ldots \rangle$ are also strictly decreasing, with limit 0.

As a result, we have that any point $x \in (0,1]$ uniquely lies in an interval of the form $(b_i, a_i]$ or of the form $(a_{i+1}, b_i]$. We can thus define a function $f : [0,1] \to \mathbb{R}$, as follows: set $f(0) = 0$, and for any $x \in (0,1]$, if $x \in (b_i, a_i]$, let $f(x)$ be such that the point $(x, f(x))$ lies on the segment connecting the point $(b_i, 0)$ to the point (a_i, a_i) (i.e., $f(x) = a_i(x - b_i)/(a_i - b_i)$); otherwise (if $x \in (a_{i+1}, b_i]$), let $f(x)$ be such that $(x, f(x))$ lies on the segment connecting (a_{i+1}, a_{i+1}) to $(b_i, 0)$. Since each of these segments clearly lies between the lines $y = x$ and $y = 0$, we have that $0 \leqslant f(x) \leqslant x$ for all $x \in [0,1]$. (*See Figure 1.*)

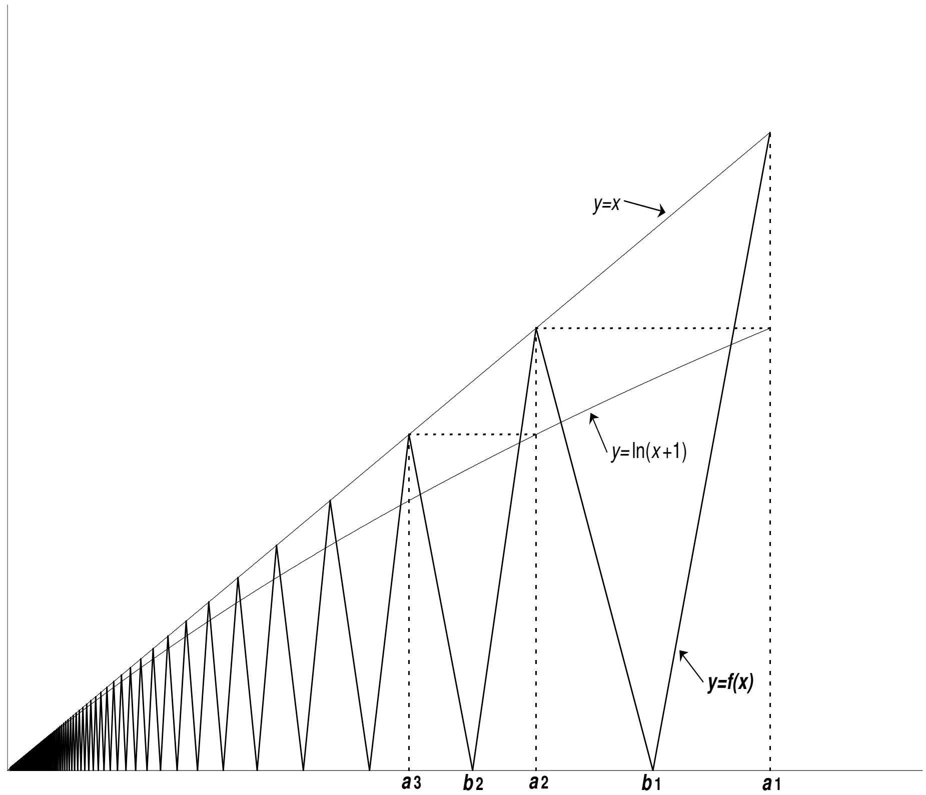

Figure 1: The construction of $y = f(x)$ over $[0, 1]$

Extend f to $[0,\infty)$ by setting $f(x)=1$ for $x>1$, and further extend f to the whole of $\mathbb{R}$ by setting $f(x)=-f(-x)$ for $x<0$. f is plainly a well-defined continuous function, and is moreover classically differentiable everywhere outside of the set $\{\pm a_i, \pm b_i \mid i \geqslant 1\} \cup \{0\}$.

Lemma 4.2 *f is not classically differentiable at $x=0$.*

Proof. This easily follows from the fact that f takes on the values (a_i, a_i) and $(b_i, 0)$ for all i. Since the sets $\{a_i \mid i \geqslant 1\}$ and $\{b_i \mid i \geqslant 1\}$ are dense at 0, the classical derivative cannot exist as it would have to simultaneously take on the values (among others) of 1 and 0. □

Lemma 4.3 *f is informatically differentiable at $x=0$ and $df_\mu(0)=1$.*

Proof. Observe first that since f is an odd function, it suffices to consider intervals of the form $[0,\delta]$ when computing the limit underlying the informatic derivative.

Let therefore $I=[0,\delta]$ be a non-trivial interval of length at most 1, and find the least a_i such that $\delta \leqslant a_i$. We then have $a_{i+1}<\delta$, so $[0,a_{i+1}] \subset I$, which entails that $\bar{f}([0,a_{i+1}]) \subseteq \bar{f}(I)$, and thus $\mu\bar{f}([0,a_{i+1}]) \leqslant \mu\bar{f}(I)$. Note that $\mu\bar{f}([0,a_{i+1}])=a_{i+1}$, since $f(a_{i+1})=a_{i+1}$. Combining, we have

$$\frac{l(a_i)}{a_i} = \frac{a_{i+1}}{a_i} \leqslant \frac{a_{i+1}}{\delta} = \frac{\mu\bar{f}([0,a_{i+1}])}{\mu I} \leqslant \frac{\mu\bar{f}(I)}{\mu I} \leqslant 1 \ .$$

If we now let δ tend to 0, we get that a_i tends to 0 as well, and hence that the ratio $l(a_i)/a_i = (l(0+a_i)-l(0))/a_i$ tends to $l'(0)=1$. Consequently,

$$d\bar{f}_\mu[0] = \lim_{I\to[0]} \frac{\mu\bar{f}(I)}{\mu I} = 1 \ .$$

We conclude the proof by noting that, clearly, $\mathrm{sgn}(d\bar{f}_\mu)(0)=1$. □

We point out that it would be easy, if slightly messy, to convert f into a function classically differentiable over $\mathbb{R}-\{0\}$: one would need to 'smooth' f over each of its cusps. The resulting function would be informatically differentiable over the whole of $\mathbb{R}$, yet still not classically differentiable at zero. We therefore record the following:

Theorem 4.4 *In what follows, p, a, and b are real numbers with $a<b$, and f is a continuous function from $\mathbb{R}$ to $\mathbb{R}$.*

1. *There are functions f which admit an informatic derivative at p yet which are not classically differentiable at p.*

2. *There are functions f which are informatically differentiable over (a,b) yet which fail to be classically differentiable over an infinite subset of (a,b).*

3. *Any function f which is informatically differentiable over (a,b) must be classically differentiable over a dense subset of (a,b).*

Proof. (Sketch.) (ii) can be achieved by iterating the construction of f (including the smoothing-over) over smaller and smaller sub-intervals of (a, b). For (iii), first note that for any $x \in (a, b)$, either $d\bar{f}_\mu$ is strictly positive in some open interval about x, or $d\bar{f}_\mu$ vanishes at points arbitrarily close to x; in either case, we can invoke Theorem 2.1 to conclude that f' exists in a set with accumulation point x. $\square$

5 Conclusion

We have shown that notions of informatic and classical differentiation are tightly related, yet that the former strictly generalizes the latter. As a consequence, it may be possible to extend techniques and results which belong to the classical differential calculus to functions that need not be classically differentiable.

For example, a standard theorem of calculus states that if $g(p) = p$ and $|g'(p)| < 1$, then the fixed point p is a local attractor for g. An entirely similar result holds in the informatic world [3, 4]. Setting $g(x) = \alpha \cdot f(x)$, where f is the function constructed in the previous section and $0 \leqslant \alpha < 1$, we are therefore able to conclude that 0 is a locally attractive fixed point of g, even though $g'(0)$ does not exist. Along the same lines, it may be possible to extend iterative zero-finding techniques such as Newton's method to functions that need not be classically differentiable.

The work presented in this paper only scratches the surface; it would be very interesting to further investigate the links between informatic and classical differentiation. In particular, it would be most useful to develop a full differential (and integral?) informatic calculus, and perhaps even port some of these ideas back over to the broader context of domain theory.

It may also be very fruitful to contrast the notions of informatic and nonsmooth differentiation [2].

References

[1] S. Abramsky and A. Jung. Domain theory. In *Handbook of logic in computer science, Vol. III.* Oxford University Press, 1994.

[2] F. H. Clarke, Y. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth analysis and control theory.* Springer-Verlag, New York, 1998.

[3] K. Martin. *A foundation for computation.* PhD thesis, Department of Mathematics, Tulane University, 2000.

[4] K. Martin. Powerdomains and zero finding. In *Proceedings of QAPL 01.* ENTCS 59(3), 2001.

[5] K. Martin. The informatic derivative at a compact element. In *Proceedings of FoSSaCS 02*, volume 2303, pages 310–325. Springer LNCS, 2002.