

# Estimation of the Derivatives of a Digital Function with a Convergent Bounded Error

Laurent Provot and Yan Gérard

Univ. Clermont 1, ISIT, Campus des Cézeaux, 63172 Aubière, France  
provot.research@gmail.com, yan.gerard@u-clermont1.fr

**Abstract.** We provide a new method to estimate the derivatives of a digital function by linear programming or other geometrical algorithms. Knowing the digitization of a real continuous function  $f$  with a resolution  $h$ , this approach provides an approximation of the  $k^{\text{th}}$  derivative  $f^{(k)}(x)$  with a maximal error in  $O(h^{\frac{1}{1+k}})$  where the constant depends on an upper bound of the absolute value of the  $(k+1)^{\text{th}}$  derivative of  $f$  in a neighborhood of  $x$ . This convergence rate  $\frac{1}{k+1}$  should be compared to the two other methods already providing such uniform convergence results, namely  $\frac{1}{3}$  from Lachaud *et. al* (only for the first order derivative) and  $(\frac{2}{3})^k$  from Malgouyres *et al.*

**Keywords:** Derivative estimation, Digital Level Layer, Convergence rate, Linear Programming.

## 1 Introduction

One of the main goal of digital geometry is to provide a theory of digital differential geometry valid on digital surfaces or digital functions. In this framework, the question of the derivative of a function at any order  $k$  is central. To be suitable for applications, we are waiting from a digital derivative that it remains close from a continuous derivative if we apply it on the restriction of the function on a digital subset around a point. More formally, it means that we want a property of uniform convergence if the resolution  $h$  tends to 0. Such a property of digital derivatives has been first investigated by Vialard, Lachaud, de Vieilleville, Feschet in [7,4,8] and Brunet, Fourey, Malgouyres, Esbelin in [5,1]. The first approach is based on the estimation of digital tangents and provides a uniform error in  $O(h^{1/3})$  for the first order derivative while the second approach dealing with binomial convolutions has a maximal error in  $O(h^{(2/3)^k})$  for the derivative of order  $k$ . The purpose of the paper is to provide a third method based on Taylor polynomial with a uniform error in  $O(h^{\frac{1}{k+1}})$ . The drawback of the method is that it depends on a parameter of maximal roughness.

The principle is to fit the values of a digital function by a polynomial. Exact fitting, namely interpolation, is not always possible but we expand the values of the function  $F(x)$  in intervals  $[F(x) - R; F(x) + R]$  so that we can find a polynomial  $P(X)$  verifying  $P(x) \in [F(x) - R; F(x) + R]$  in a neighborhood of  $x_0$ . How can we choose  $R$ ? and the neighborhood of  $x_0$ ?

We start the paper by introducing the notion of roughness of a function, namely the minimum  $R$  necessary to find a polynomial approximating the function  $F$  with an error smaller than  $R$ . We then use this polynomial approximation  $P$  to provide derivatives of  $F$ . In Section 4, we present theoretical results by bounding the error and in Section 5, we provide some experimental results.

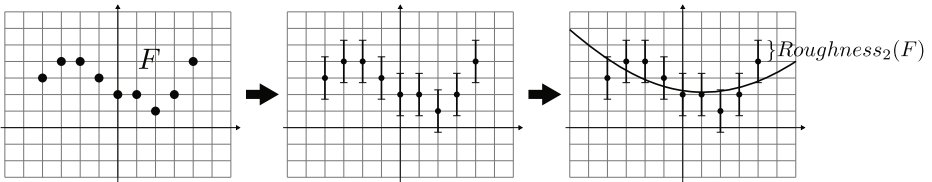
## 2 Roughness of Order $k$

### 2.1 Definition

Let  $F : X \rightarrow \mathbb{R}$  be a real function defined on a domain  $X$  which can be any subset of  $\mathbb{R}$ , but most often a real interval,  $\mathbb{Z}$  or a finite subset of consecutive integers. A classic idea about any function  $f$  is to approximate its values by a polynomial  $P(X) \in \mathbb{R}^k[X]$  of bounded degree. Except in degenerated cases or if the cardinality of  $X$  is less than or equal to  $k + 1$ , it is not possible to find a polynomial in  $\mathbb{R}^k[X]$  fitting exactly  $F$ , namely satisfying for all  $x$  in  $X$  the equalities  $F(x) = P(x)$ . Hence approximating methods have been developed, such as Least Squares Fitting. In the framework of this paper, we focus on uniform fitting:

**Definition 1.** For any function  $F : X \rightarrow \mathbb{R}$  with  $X \subset \mathbb{R}$  and any fixed order  $k$ , the roughness of  $F$  of order  $k$ , denoted  $Roughness_k(F)$  is the lower bound of the values  $R$  such that there exists a polynomial  $P(X)$  in  $\mathbb{R}^k[X]$  fitting  $F$  with a uniform error smaller than  $R$ , namely

$$Roughness_k(F) = \inf_{P \in \mathbb{R}^k[X]} \left( \sup_{x \in X} |P(x) - F(x)| \right).$$



**Fig. 1.** On the left, a function  $F$ . In the middle, we expand each value of  $F$  in an interval as small as possible so that there exists a polynomial of degree at most 2 passing through each vertical segment (drawn on the right). This minimal radius around the values of  $F$  is the roughness of order 2 of  $F$  since we use here polynomials of degree at most 2.

What does it mean in practice? That instead of considering the exact value  $F(x)$  at each  $x$  of  $X$ , we expand it in an interval  $I_x = [F(x) - R; F(x) + R]$ . There is a threshold  $R_k(F)$  such that, with  $R < R_k(F)$ , there exists no polynomial of degree at most  $k$  verifying  $P(x) \in I_x$  for any  $x \in X$  while for  $R > R_k(F)$  there exists at least one polynomial satisfying these conditions (Fig. 1). This polynomial is used in next sections to provide an estimation of the derivative of order  $k$  of  $F$ .

We can also notice that if  $X$  is finite, the lower bound becomes a minimum. If  $X$  is not bounded, it is possible to find an infinite roughness, for instance if  $F$  is an exponential, its difference with a polynomial can not be bounded. Last, we add that the sequence  $Roughness_k(F)$  is decreasing with  $k$  and decreases until 0 if  $X$  is finite. We have  $R_k(F) = 0$  when  $k \geq |X| - 1$  (a null roughness means that there exists an exact fitting of  $F$  by a polynomial).

### 2.2 Taylor-Lagrange Inequality

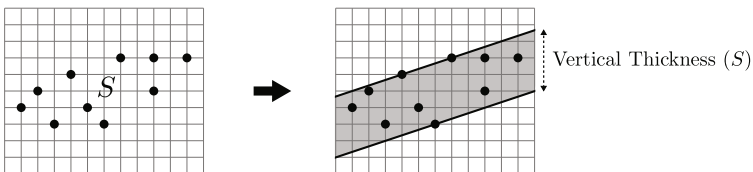
Let us recall Taylor-Lagrange inequality for a function  $F : \mathbb{R} \rightarrow \mathbb{R}$  of class  $C^{k+1}$ . We consider the Taylor Polynomial  $T_{x_0}(x) = \sum_{i=0}^k \frac{f^{(i)}(x_0)}{i!} (x - x_0)^i$  of degree  $k$ . If we assume that  $|f^{(k+1)}(x)| \leq M$  for  $x \in [x_0 - r; x_0 + r]$ , we have the inequality  $|F(x) - T_{x_0}(x)| \leq M \frac{(x-x_0)^{k+1}}{(k+1)!}$ . It means that for any  $x$  in the interval  $[x_0 - r; x_0 + r]$ , the uniform distance between the Taylor polynomial at  $x_0$  and  $F$  is less than or equal to  $M \frac{r^{k+1}}{(k+1)!}$ . By taking  $x_0 = \frac{a+b}{2}$  and  $2r = b - a$ , it leads to the next property:

*Property 1.* The roughness of the restriction  $F_{[a;b]}$  of a function  $F$  on an interval  $[a; b]$  is smaller than  $\|F_{[a;b]}^{(k+1)}\|_{\infty} \frac{(\frac{b-a}{2})^{k+1}}{(k+1)!}$ .

### 2.3 Roughness and Vertical Thickness

The roughness of order 1 is related to the vertical thickness of the graph of  $F$  but it is not only true for the order 1. Let us recall the definition of the thickness: Given a finite subset  $S$  of  $\mathbb{R}^d$  (we call *vertical* the direction of the last coordinate), we define its vertical thickness as the minimal value  $\delta$  such that all points  $x$  of  $S$  verify  $h \leq n.x \leq h + \delta$  where the last coordinate of the normal vector  $n$  is 1. The double inequality  $h \leq n.x \leq h + \delta$  simply means that the point  $x$  is between the hyperplane of equation  $h = n.x$  and its translation by vector  $(0, \dots, 0, \delta)$ . Hence the vertical thickness is the minimal vertical height of a strip containing  $S$  between two parallel hyperplanes (Fig. 2).

Now if we look at the double inequalities defining the roughness of order  $k$  of  $F$ , we find exactly the same kind of constraints: By denoting  $a_i$  the coefficients of polynomial  $P(X)$ , the double inequality  $F(x) - R \leq P(x) \leq F(x) + R$  can be rewritten  $a_0 - R \leq -\sum_{i=1}^k a_i x^i + 1F(x) \leq a_0 + R$ . This is exactly  $h \leq n.x \leq h + \delta$



**Fig. 2.** On the left, a finite set  $S$ . On the right, the vertical thickness of  $S$  is the minimal height of a strip containing  $S$ .

where  $h$  plays the role of  $a_0 - R$ , where  $\delta$  is equal to  $2R$ , where the unknown normal vector  $n$  (of last coordinate 1) corresponds to the opposite of the unknown coefficients of the polynomial completed by a last coordinate equal to 1 namely  $n = ((-a_i)_{1 \leq i \leq k}; 1)$  and where the input points  $x$  are given by  $((x^i)_{1 \leq i \leq k}; F(x))$  for all  $x$  in  $X$ . It leads to the property:

*Property 2.* The roughness of order  $k$  of  $F$  is the half of the vertical thickness of the set  $S = \{((x^i)_{1 \leq i \leq k}; F(x)) \text{ for } x \in X\}$ .

## 2.4 Computation

The roughness can be computed by Linear Programming or directly by some algorithms of Computational Geometry.

**Linear Programming.** For computations, we consider a function  $F : X \rightarrow \mathbb{R}$  defined on a finite domain  $X$ . By denoting  $a_i$  the coefficients of the fitting polynomial  $P(X) = \sum_{i=1}^k a_i X^i$ , the roughness of order  $k$  of  $F$  is the minimum of the objective function  $R$  under linear constraints  $F(x) - R \leq \sum_{i=1}^k a_i x^i \leq F(x) + R$  for any  $x \in X$  (variables are coefficients  $a_i$  and  $R$  while the values  $x$  and  $F(x)$  are given in the input). Hence the minimum of  $R$  can be computed by any algorithm of linear programming (Simplex, Interior Points, ...). It means that its value can be obtained in linear time in the size of  $X$  since linear programming in fixed dimension can be solved in  $O(n)$  time where  $n$  is the number of linear constraints [6].

**Computational Geometry.** According to property 2, we can compute the roughness as the vertical thickness of a set in dimension  $k + 1$ . The first idea to compute this thickness is to pass through the convex hull of the set of points. There is however another faster approach based on the chords set: the vertical thickness of the set  $S$  can be obtained as the height of the facet of the convex hull of  $S - S$  (the chords) cutting the vertical axis (see [2] for details). The principle of the chords algorithm is to climb all along the vertical axis with the same kind of ideas used in QuickHull or GJK [3]. The time of computation is quasi-linear in practice but its generalization in  $nD$  requires new ideas in order to avoid loops.

## 3 Digital Estimation of $F^{(k)}(0)$

### 3.1 Parameter

Our estimation of the order  $k$  derivative of a function  $F$  defined on a discrete domain  $X$  at a given point is a direct corollary of the previous notion of roughness. For convenience, we consider a function  $F : \mathbb{Z} \rightarrow \mathbb{Z}$ . The aim of this section is to provide a method to estimate a kind of  $k^{th}$  derivative of  $F$  at 0. By translation, it provides a method to estimate the derivatives  $F^{(k)}(x)$  at any integer. Unfortunately, we do not provide a canonical definition of these values. The estimation method we provide depends on a parameter of maximal roughness

$R_{max}$ . We will see in the next section that by choosing this parameter greater than  $\frac{1}{2} + \frac{\|f^{(k+1)}(x)\|_\infty}{(k+1)!}$ , we obtain nice results of convergence of the digital derivative of maximal roughness  $R_{max}$  towards the real continuous derivative.

It leads of course to the question: can the maximal roughness be chosen automatically? Previous remark relates it to an upper bound of the  $k + 1$ -derivative around the considered point. Hence a possible maximal roughness  $R_{max}$  can be estimated with finite differences (for instance  $\frac{F(1)-F(-1)}{2}$  for an estimation of  $F'(0)$ ,  $F(1) + F(-1) - 2F(0)$  for an estimation of  $F''(0)$ , ...). As we will see in experiments, another approach is to fix *a priori* this threshold between  $\frac{1}{2}$  and 1 ( $\frac{1}{2}$  seems to be a minimum to take into account the gaps of the digitization).

### 3.2 Definition

We consider the restriction  $F_{[-m;m]}$  of the function  $F$  in a neighborhood of 0 going from  $-m$  to  $m$ . Given the maximal roughness parameter  $R_{max}$ , if  $m$  is small enough, the roughness of order  $k$  of  $F_{[-m;m]}$  is smaller than this maximum authorized  $R_{max}$  (if  $m = 0$ , the roughness at any order  $k$  is null). The idea is to increase  $m$  until the roughness of  $F_{[-m;m]}$  becomes greater than the fixed parameter  $R_{max}$ . The intervals  $[-m; m]$  providing a roughness smaller than the limit are called  $k$ -neighborhood of 0 wrt.  $F$ :

**Definition 2.** *Given  $F : \mathbb{Z} \rightarrow \mathbb{Z}$  and a maximal roughness  $R_{max}$ , a  $k$ -neighborhood of 0 is an integer interval  $[-m; m]$  such that the restriction  $F_{[-m;m]}$  has a roughness of order  $k$  less than  $R_{max}$ . The maximal  $k$ -neighborhood  $[-m_{max}; m_{max}]$  of 0 is the largest  $k$ -neighborhood.*

There are two versions of these definitions, one with a strict inequality  $Roughness_k(F_{[-m;m]}) < R_{max}$  and another one where we allow a large inequality  $Roughness_k(F_{[-m;m]}) \leq R_{max}$ . At this step, we are not enough advanced to determine if one is better than the other. We can just notice that if we want the notion of maximal neighborhood of order 1 to collapse with the notion of digital tangent already defined in digital geometry [7], we should choose  $R_{max} = \frac{1}{2}$  and a strict inequality since digital straight segments have a vertical thickness strictly less than 1. This leads to the introduction of the derivative of order  $k$ :

**Definition 3.** *Let  $[-m; m]$  be the maximal  $k$ -neighborhood of 0 for a function  $F$  of maximal roughness  $R_{max}$ . There exists at least a polynomial  $P(X) = \sum_{i=1}^k a_i x^i$  in  $\mathbb{R}^k[X]$  verifying  $F(x) - Roughness_k(F_{[-m;m]}) \leq P(x) \leq F(x) + Roughness_k(F_{[-m;m]})$  for any integer  $x$  from  $-m$  to  $m$ . We define the  $k^{th}$  derivative of  $F$  at 0 by  $F^{(k)}(0) = k!a_k$ .*

Attentive readers will notice that it can arise non unique solutions  $P(X)$  in degenerated cases. In this case, which coefficient  $a_k$  should we choose to define the derivative? Again, depending on the application we can leave open the choice, consider that there is no good value and that it is better to provide an interval as a solution, or that the mean on the interval  $[a_{k_{min}}; a_{k_{max}}]$  is the best representative of the possible derivatives that can be chosen.

Of course, this definition does not match all the properties of continuous derivative. We have for instance no guarantee that by derivating a derivative of order  $k$ , we obtain the same value as by computing directly the derivative of order  $k+1$ . Hence how can we say that this definition is suitable? with a property we investigate in the next section in order to show that the digital derivative estimation converges towards the continuous value.

### 4 Error Bounding

Let us consider the framework of Numerical Analysis or Signal Theory. We consider a real function  $f : \mathbb{R} \rightarrow \mathbb{R}$  of class  $C^{k+1}$  known from its digitization in a grid of horizontal and vertical resolution  $h$ . It means that we know a portion of the digital function  $F : \mathbb{Z} \rightarrow \mathbb{Z}$  defined by  $F(x) = \lfloor \frac{f(hx)}{h} \rfloor$ . We can estimate the derivatives of  $F$  around 0 with a given maximal roughness by the method described in the previous section. As we have here the initial value of  $f$ , the question is to determine the maximal error of this method of estimation on the real derivative of order  $k$  –corrected by a scale factor  $h^{1-k}$ – namely to bound the difference  $f^{(k)}(0) - h^{1-k}F^{(k)}(0)$ . This scale factor comes from the fact that the derivative of order  $k$  of  $x \mapsto \frac{f(hx)}{h}$  is  $(\frac{f(hx)}{h})^{(k)} = h^{k-1}f^{(k)}(hx)$ . We provide the following upper bound:

**Theorem 1.** *Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be a real function of class  $C^{k+1}$  whose derivative of order  $k + 1$  is bounded by  $M$  in a neighborhood of 0 (more precisely we assume  $|f^{(k+1)}(x)| \leq M$  for  $x \in [-L; L]$ ). We have  $|f^{(k)}(0) - h^{1-k}F^{(k)}(0)| \leq u_k(\frac{1}{2} + \frac{M}{(k+1)!} + R_{max})h^{\frac{1}{k+1}}$  for the estimation of  $F^{(k)}(0)$  with roughness  $R_{max} \geq \frac{1}{2} + \frac{M}{(k+1)!}$ .*

The proof relies on two inequalities: the first one says that the difference between Taylor polynomial  $T(x)$  of  $\frac{f(hx)}{h}$  around 0 of order  $k$  and  $F(x) = \lfloor \frac{f(hx)}{h} \rfloor$  is small. The second one expresses the fact that the difference between  $\lfloor \frac{f(hx)}{h} \rfloor$  and the polynomial  $P(x)$  that coefficients are obtained by computations of the approximation  $F$  with maximal roughness  $R_{max}$  is also small in a discrete neighborhood of 0. It leads to bound the difference between  $P(x)$  and  $T(x)$  ( $\frac{f(hx)}{h}$  plays the role of intermediary value) on a discrete neighborhood. With an important lemma on discrete norms on polynomials, we obtain the claimed result. Details of the proof are given in next subsections.

#### 4.1 Bounding the Difference $T(x) - P(x)$ on a Discrete Neighborhood

It is the first step of the proof of Theorem 1. We consider values of  $x$  in the interval  $[-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}]$  with  $h$  small enough to have  $hx$  included in  $[-L; L]$  for all these values (in other words  $h$  is chosen smaller than  $L^{k+1}$ ). It follows that the  $(k + 1)^{th}$  derivative of  $\frac{f(hx)}{h}$  namely  $h^k f^{(k+1)}(hx)$  verifies  $(\frac{f(hx)}{h})^{(k+1)} \leq h^k M$ .

According to Taylor-Lagrange inequality at point 0 for function  $\frac{f(hx)}{h}$ , we have  $|T(x) - \frac{f(hx)}{h}| \leq \frac{h^k M}{(k+1)!} x^{k+1} \leq \frac{h^k M}{(k+1)!} (h^{-\frac{k}{k+1}})^{k+1}$  namely

$$\left| T(x) - \frac{f(hx)}{h} \right| \leq \frac{M}{(k+1)!} \tag{1}$$

where  $T(x)$  is the Taylor polynomial of  $\frac{f(hx)}{h}$  at 0:  $T(x) = f(0) + \sum_{i=1}^k h^{i-1} \frac{f^{(i)}(0)}{i!} x^i$ . As we have  $0 \leq \frac{f(hx)}{h} - \lfloor \frac{f(hx)}{h} \rfloor \leq 1$ , it follows that  $\left| \lfloor \frac{f(hx)}{h} \rfloor + \frac{1}{2} - \frac{f(hx)}{h} \right| \leq \frac{1}{2}$  and with bounding (1):

$$\forall x \in \left[ -h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}} \right], \quad \left| T(x) - \frac{1}{2} - \left\lfloor \frac{f(hx)}{h} \right\rfloor \right| \leq \frac{1}{2} + \frac{M}{(k+1)!} \tag{2}$$

Now we can be interested in the bounding of the difference between  $P(x)$  and  $F(x) = \lfloor \frac{f(hx)}{h} \rfloor$ . According to inequality (2) and condition  $R_{max} \geq \frac{1}{2} + \frac{M}{(k+1)!}$  of Theorem 1, it follows that the interval  $[-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}]$  is a neighborhood of 0 for maximal roughness  $R_{max}$ . Hence the maximal neighborhood of 0 for the function  $F$  is at least the interval  $[-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}]$ .

Now let us introduce the polynomial  $P(x)$  obtained with our method to estimate the derivative of a digital function : it is a solution of inequalities  $|P(x) - F(x)| \leq Roughness_k(F_{[-m;m]})$  where  $[-m;m]$  is the maximal  $k$ -neighborhood of 0 for the function  $F$  and maximal roughness  $R_{max}$ . As the maximal neighborhood contains  $[-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}]$ , we have:

$$\forall x \in [-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}] \cap \mathbb{Z}, \quad \left| P(x) - \left\lfloor \frac{f(hx)}{h} \right\rfloor \right| \leq R_{max} \tag{3}$$

It follows directly from inequalities (2) and (3)

$$\forall x \in [-h^{-\frac{k}{k+1}}; h^{-\frac{k}{k+1}}] \cap \mathbb{Z}, \quad \left| T(x) - \frac{1}{2} - P(x) \right| \leq \frac{1}{2} + \frac{M}{(k+1)!} + R_{max} \tag{4}$$

### 4.2 Discrete Norms on Polynomials

The aim of this subsection is to show that two polynomials whose difference can be bounded on a discrete interval  $\{-m\dots m\}$  – as previous polynomials  $T - \frac{1}{2}$  and  $P$  – have necessarily close coefficients.

**Lemma 1.** *We assume  $k < 2m$ . Let  $A(X)$  and  $B(X)$  be two polynomials of degree at most  $k$  whose difference  $|A(X) - B(X)|$  is less than a constant  $C$  on a discrete interval  $\{-m\dots m\}$ . We have a difference  $|a_k - b_k|$  of coefficients of degree  $k$  which is smaller than  $Cv_k m^{-k}$  where  $v_k$  is a constant depending only on  $k$ .*

Let us denote  $\alpha_k(m, n)$  the maximum of the coefficient of degree  $k$  among all polynomials  $P(X)$  of degree at most  $n$  and verifying  $|P(x)| \leq 1$  for all integers  $x$  from  $-m$  to  $m$ .

*Property 3.* – The value  $\alpha_k(m, n)$  decreases with  $m$ .

- The value  $\alpha_k(m, n)$  increases with  $n$ .
- For any positive integer  $\mu$ , we have  $\alpha_k(\mu m, n) \leq \alpha_k(m, n)/\mu^k$ .
- We have  $1/m^k \leq \alpha_k(m, n)$
- If we assume  $m \geq 2p$  where  $p = \lfloor (n - 1)/2 \rfloor + 1$ , we have  $\alpha_k(m, n) \leq (2p)^k \alpha_k(p, n)/m^k$  namely  $m^k \alpha_k(m, n)$  is bounded as a function of  $m$ .

*Proof.* The decrease of  $\alpha_k(m, n)$  in  $m$  and increase in  $n$  are straightforward.

For going from  $\alpha_k(\mu m, n)$  to  $\alpha_k(m, n)$ , we just have to notice that if  $P(X)$  satisfies  $-1 \leq P(x) \leq 1$  for any integer  $x$  from  $-\mu m$  to  $\mu m$ , then  $P(\mu X)$  whose coefficient of degree  $k$  is  $\mu^k a_k$  verifies  $-1 \leq P(\mu x) \leq 1$  for all integers  $x$  from  $-m$  to  $m$ . It follows that  $|\mu^k a_k| \leq \alpha_k(m, n)$ . It proves that the maximum of  $|a_k|$  is at most  $\alpha_k(m, n)/\mu^k$  namely  $\alpha_k(\mu m, n) \leq \alpha_k(m, n)/\mu^k$ .

To prove  $1/m^k \leq \alpha_k(m, n)$ , we just have to notice that the polynomial  $P(X) = X^k/m^k$  verifies  $-1 \leq P(x) \leq 1$  for all integer  $x$  from  $-m$  to  $m$ . Hence the maximum  $\alpha_k(m, n)$  of the coefficient of degree  $k$  under these constraints is at least  $1/m^k$ .

For the last inequality, we define  $p = \lfloor (n - 1)/2 \rfloor + 1$  because it is the first integer for which  $\alpha_k(p, n)$  is defined (it verifies  $2p + 1 > n$ ). The assumption  $m \geq 2p$  allows to write  $m \leq 2m - 2p$  which leads to (i)  $m/p \leq 2(m - p)/p = 2(m/p - 1)$ . As the function  $\alpha_k(m, n)$  increases with  $m$ , we have  $\alpha_k(m, n) \leq \alpha_k(\lfloor m/p \rfloor p, n)$ . With  $\mu = \lfloor m/p \rfloor$ , it provides (ii)  $\alpha_k(m, n) \leq \alpha_k(p, n)/\lfloor m/p \rfloor^k$ . If we look at the denominator  $\lfloor m/p \rfloor^k$ , we have  $\lfloor m/p \rfloor > m/p - 1 \geq m/2p$  according to (i). Hence we obtain  $\lfloor m/p \rfloor^k > (m/2p)^k$ . In our main inequality (ii), it leads to  $\alpha_k(m, n) \leq (2p)^k \alpha_k(p, n)/m^k$ .

The property 3 shows that the values  $\alpha_k(m, n)$  decrease in  $1/m^k$ . Lemma 1 is just a consequence of this property for  $n = k$ , i.e  $\alpha_k(m, k) \leq \frac{v_k}{m^k}$ .

### 4.3 Some Values for $\alpha_k(m, n)$

- With  $n = 0$ , we have  $\alpha_0(m, 0) = 1$ . More generally we have  $\alpha_0(m, n) = 1$ .
- With  $n = 1$ , we have  $\alpha_1(m, 1) = 1/m$ .
- With  $n = 2$ , we have  $\alpha_1(m, 2) = 1/m$  and  $\alpha_2(m, 2) = 2/m^2$ .
- With  $n = 3$ , we have  $\alpha_1(m, 3) = \frac{3}{m}$  if  $m$  is even and  $\alpha_1(m, 3) = \frac{3m^2 + 1}{(m - 1)m(m + 1)}$  if  $m$  is odd.  $\alpha_2(m, 3) = 2/m^2$ .  $\alpha_3(m, 3) = \frac{4}{m^3}$  if  $m$  is even and  $\alpha_3(m, 3) = \frac{4}{(m - 1)m(m + 1)}$  if  $m$  is odd.
- With  $n = 4$ , we have  $\alpha_1(m, 4) = \alpha_1(m, 3)$ ,  $\alpha_2(m, 4) = \frac{2m^2}{(m - i)(m + i)i^2}$ ,  $\alpha_3(m, 4) = \alpha_3(m, 3)$  and  $\alpha_4(m, 4) = \frac{2}{(m - i)(m + i)i^2}$  where  $i$  is the closest integer from  $\frac{m}{\sqrt{2}}$ .

### 4.4 Synthesis

To obtain Theorem 1, we just apply Lemma 1 on inequality (3) with  $T(X) - \frac{1}{2}$  and  $P(X)$  as polynomials  $A(X)$  and  $B(X)$ ,  $\frac{1}{2} + \frac{M}{(k + 1)!} + R_{max}$  as constant  $C$  and  $m = \lfloor h^{-\frac{k}{k + 1}} \rfloor$ . We first recall that our estimation of the derivative  $F^{(k)}(0)$  is  $k!a_k$



where  $a_k$  is the coefficient of degree  $k$  of  $P$ . For  $k > 0$ , the coefficient of degree  $k$  of  $T(X) - \frac{1}{2}$  is  $h^{k-1} \frac{f^{(k)}(0)}{k!}$ . It follows  $\left| \frac{F^{(k)}(0)}{k!} - h^{k-1} \frac{f^{(k)}(0)}{k!} \right| \leq \left( \frac{1}{2} + \frac{M}{(k+1)!} + R_{max} \right) \frac{v_k}{m^k}$  namely  $\left| h^{1-k} \frac{F^{(k)}(0)}{k!} - f^{(k)}(0) \right| \leq \left( \frac{1}{2} + \frac{M}{(k+1)!} + R_{max} \right) k! v_k \frac{h^{1-k}}{m^k}$ . It remains to say that for  $h$  small enough (according to  $k$ ) we have  $m = \lfloor h^{-\frac{k}{k+1}} \rfloor \geq (1 - \epsilon) h^{-\frac{k}{k+1}}$  and thus  $m^{-k} \leq (1 - \epsilon)^{-k} m^{\frac{k^2}{k+1}}$ . It leads to  $\left| h^{1-k} \frac{F^{(k)}(0)}{k!} - f^{(k)}(0) \right| \leq \left( \frac{1}{2} + \frac{M}{(k+1)!} + R_{max} \right) k! (1 - \epsilon)^{-k} v_k h^{1-k + \frac{k^2}{k+1}}$ . With  $u_k = k! (1 - \epsilon)^{-k} v_k$  it provides Theorem 1.

### 4.5 State of the Art

The previous approach based on Taylor Polynomial provides a numerical algorithm to estimate the derivatives of a function known from its digitization with a guaranteed result. It is however not the first method to solve this problem. Several works have been developed in order to estimate the derivatives of such a function. Lachaud *et al.* provide a method with a mean convergence rate in  $O(h^{2/3})$  and a worst case convergence rate in  $O(h^{1/3})$  to estimate the tangent of the curve, namely its first order derivative [4]. This approach is based on maximal digital segment recognition and is thus rather close to the approach provided here (the difference is in the choice of the maximal thickness or roughness). The construction of a curvature estimator that has the property of asymptotic convergence was however still open in this framework [8]. More recently, Malgouyres *et al.* provide a new method based on binomial estimation allowing to estimate the derivatives at any order [5,1]. The maximal error of this  $k^{th}$  derivative estimator converges in  $O(h^{(2/3)^k})$ .

It makes at least two other methods to estimate the derivatives of a function known from its digitization with a controlled uniform error. As far as we know, no other result has been published in this specific framework. Hence there are three methods to compare according to at least three theoretical criteria:

**The speed of convergence:** The speeds of convergence of the three methods are of the form  $O(h^d)$  where  $h$  is the resolution (tending to 0) and  $d$  is the degree. The higher the exponent  $d$ , the better the convergence. For the method based on discrete segments, we have a degree (in the worst case) of  $\frac{1}{3}$  only for the first order derivative. For the binomial method the degree of convergence is  $(\frac{2}{3})^k$  to estimate the derivative of order  $k$  while the approach based on Taylor Polynomial and Linear Programming provides a degree of  $\frac{1}{k+1}$ . We compare these three degrees for the orders of derivatives from 1 to 5 (Tab. 1).

**Table 1.** Comparison between the degree of convergence of the three methods to estimate the  $k^{th}$  derivative

| $k$               | 1             | 2             | 3              | 4                 | 5                 |
|-------------------|---------------|---------------|----------------|-------------------|-------------------|
|                   | $1/3 = 0.333$ |               |                |                   |                   |
| $(\frac{2}{3})^k$ | $2/3 = 0.666$ | $4/9 = 0.444$ | $8/27 = 0.296$ | $16/81 = 0.19753$ | $32/243 = 0.1317$ |
| $\frac{1}{k+1}$   | $1/2 = 0.5$   | $1/3 = 0.333$ | $1/4 = 0.25$   | $1/5 = 0.2$       | $1/6 = 0.1666$    |

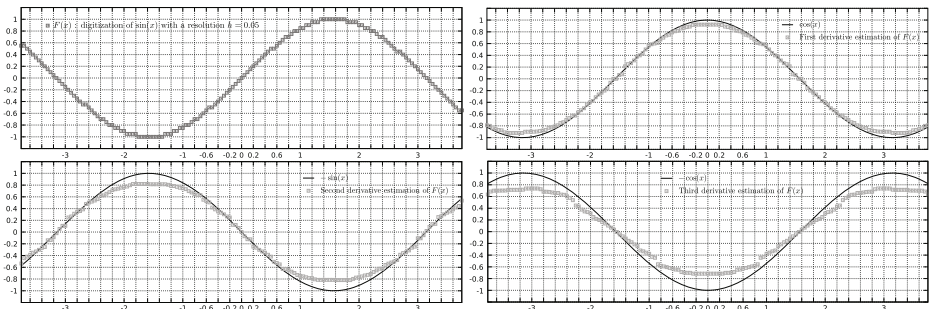
**The number of points necessary to compute the estimation:** For the approach based on maximal digital segments, the number of points is the size of the maximal segment which has no upper bound (there is an upper bound for the smallest digital segment in  $\Omega(m^{1/3} \log m)$  in [8]). With the convolution method, the first value given for the size of the mask is  $m = h^{-\frac{1}{1.01}}$  but a proposition allows to work with only  $O(\sqrt{m \ln m})$  points [1]. With the Taylor Polynomial approach, the relation  $h = m^{-1-\frac{1}{k}}$  leads to use at least  $2m + 1 = 2h^{-\frac{k}{k+1}} + 1$  points. This number of points depends on the order of the derivative. For  $k = 1$ , the degree is  $-\frac{1}{2}$ . For  $k = 2$ , it is  $-\frac{2}{3}$  and more generally  $-\frac{k}{k+1}$  for the derivative of order  $k$  (we need more points if we increase  $k$ ).

**The complexity of the computation:** In both cases, the computation time is directly related to the number of points, i.e. to the previous criterion. For the computation of the maximal digital segment, this can be done in linear time in the number of points. For the convolution, the time of computation is again linear in the size of the binomial mask  $m$ . For the Taylor Polynomial, the method uses a linear program. Hence the time of computation could be linear in  $m$  ( $O(m)$  with Megiddo algorithm [6]). But without any information on the minimal neighborhood that can be chosen to provide an approximation, we have to increase the neighborhood at each step and check that the maximal roughness is not exceeded. This process can be improved – as in the framework of digital tangents – but this naive approach provides a time of computation quadratic in the size of the maximal neighborhood.

We have compared the three approaches according to theoretical criteria but it would be probably more enlightening to compare them on a benchmark of functions whose derivative are known.

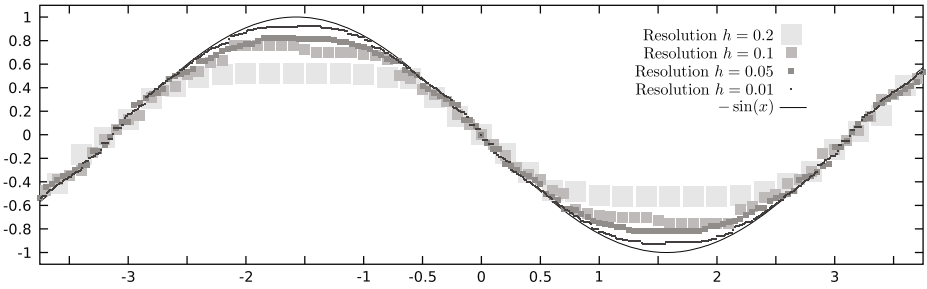
### 5 Experiments

We propose in this section some results obtained with our Taylor based method to estimate derivatives of a digital function  $F(x)$ . Fig. 3 shows a digitization

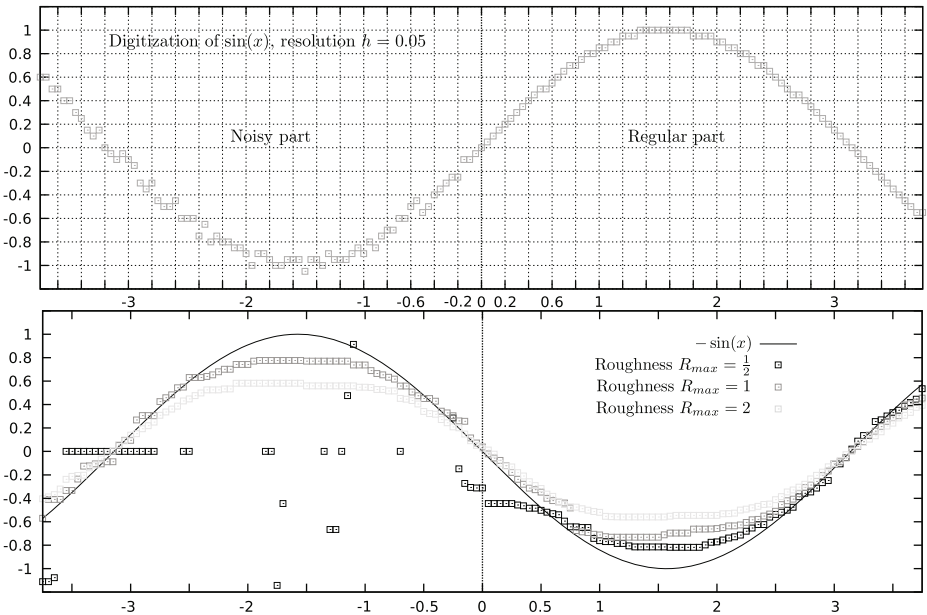


**Fig. 3.** A digitization  $F(x)$  of the sine function at resolution  $h = 0.05$  and an estimation of its derivatives from order 1 to 3 by using our Taylor based method

of the sine function (according to the grid intersect quantization (GIQ), i.e.  $F(x) = \lfloor \frac{\sin(hx)}{h} + \frac{1}{2} \rfloor$ ) with a resolution  $h = 0.05$  and an estimation of its first, second and third derivatives with a roughness  $R_{max} = \frac{1}{2}$ . The continuous derivative has been drawn for each estimation to compare with the expected theoretical results. To emphasize the influence of the resolution, the sine function has been digitized (GIQ) with different values of  $h$  and an estimation of the second derivative of this digitization (with  $R_{max} = \frac{1}{2}$ ) has been done. The results are depicted in Fig. 4. As we might expect the finer the resolution is, the more



**Fig. 4.** Estimations of the second derivative of the digitization of sine at different resolutions



**Fig. 5.** Estimations of the second derivative of the digitization of sine at different roughness. The left part ( $x < 0$ ) of the digitization is noisy to see the influence of the roughness parameter.

accurate the estimation is. Last, an estimation of the second derivative of a digitization of the sine with a resolution  $h = 0.05$  has been computed at different roughness (Fig. 5). To see the effect of this parameter, the digitization has been partially perturbed (when  $x < 0$ ) with a uniform noise. Although the most accurate results on the regular part of the digitization are obtained with the smallest roughness  $R_{max} = \frac{1}{2}$ , on the noisy part the estimation is completely distorted. If we increase the roughness the estimation becomes reliable in the noisy part, but higher roughness values tend to smooth the results.

## 6 Conclusion

We have proposed in this paper a method to estimate the derivatives of a digital function. It is a kind of generalization of the method based on digital tangents. The main idea was to relax the thickness constraint, our roughness parameter, and to allow the degree of the fitting polynomial to be greater than 1. This led to an interesting theoretical result of convergence in  $O(h^{\frac{1}{k+1}})$ . A perspective would be to extend this framework to multivariable functions but much more experiments are required to determine the relative benefits or drawbacks of this method, in terms of time of computations and accuracy of the results, compared to the convolutions approach.

## References

1. Esbelin, H.-A., Malgouyres, R.: Convergence of binomial-based derivative estimation for  $C^2$  noisy discretized curves. In: Brlek, S., Reutenauer, C., Provençal, X. (eds.) DGCI 2009. LNCS, vol. 5810, pp. 57–66. Springer, Heidelberg (2009)
2. Gerard, Y., Debled-Rennesson, I., Zimmermann, P.: An elementary digital plane recognition algorithm. *Discrete Applied Mathematics* 151, 169–183 (2005)
3. Gilbert, E.G., Johnson, D.W., Keerthi, S.S.: A fast procedure for computing the distance between complex objects in three-dimensional space. *IEEE Journal of Robotics and Automation* 4, 193–203 (1988)
4. Lachaud, J.O., Vialard, A., de Vieilleville, F.: Fast, accurate and convergent tangent estimation on digital contours. *Image and Vision Computing* 25(10), 1572–1587 (2007)
5. Malgouyres, R., Brunet, F., Fourey, S.: Binomial convolutions and derivatives estimation from noisy discretizations. In: Coeurjolly, D., Sivignon, I., Tougne, L., Dupont, F. (eds.) DGCI 2008. LNCS, vol. 4992, pp. 370–379. Springer, Heidelberg (2008)
6. Megiddo, N.: Linear programming in linear time when the dimension is fixed. *Journal of the ACM* 31(1), 114–127 (1984)
7. Vialard, A.: Geometrical parameters extraction from discrete paths. In: Miguet, S., Ubéda, S., Montanvert, A. (eds.) DGCI 1996. LNCS, vol. 1176, pp. 24–35. Springer, Heidelberg (1996)
8. de Vieilleville, F., Lachaud, J.O., Feschet, F.: Maximal digital straight segments and convergence of discrete geometric estimators. In: Kalviainen, H., Parkkinen, J., Kaarna, A. (eds.) SCIA 2005. LNCS, vol. 3540, pp. 988–997. Springer, Heidelberg (2005)