



HAL
open science

Sparse convolution-based digital derivatives, fast estimation for noisy signals and approximation results

Henri-Alex Esbelin, Rémy Malgouyres

► **To cite this version:**

Henri-Alex Esbelin, Rémy Malgouyres. Sparse convolution-based digital derivatives, fast estimation for noisy signals and approximation results. *Theoretical Computer Science*, 2016, 624, pp.2-24. 10.1016/j.tcs.2015.12.018 . hal-01296759

HAL Id: hal-01296759

<https://hal.science/hal-01296759v1>

Submitted on 1 Apr 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Sparse Convolution-based Digital Derivatives, Fast Estimation For Noisy Signals and Approximation Results

Henri-Alex Esbelin, Rémy Malgouyres

{alex.esbelin@univ-bpclermont.fr, remy.malgouyres@udamail.fr}

LIMOS UMR 6158 CNRS, Clermont-Université¹

Abstract

We provide a general notion of a Digital Derivative in 1–dimensional grids, which has real or integer-only versions. From any such masks, a family of masks called skipping masks are defined. We prove general results of multigrid convergence for skipping masks. We propose a few examples of digital derivative masks, including the now well-known binomial mask. The corresponding skipping masks automatically have multigrid convergence properties. We study the cases of parametric curves tangents and curvature. We propose a novel interpretation of digital convolutions as computing points on a smooth curve, the regularity of which depends on the mask. We establish, in the case of binomial and B –spline masks, a close relationship between the derivatives of the smooth curve, and the digital derivatives provided by the mask.

Key Words. Digital Analysis, Convolution, Digital Derivative, Derivative Estimation, Curvature Estimation

1 Introduction

1.1 Aim of the work

The goal of this work is to put in place a foundation stone for a new paradigm for numerical calculations. The motivation for this is that the objects of current methods of numerical analysis, which deal mathematically primarily with real numbers, cannot effectively be computed, or even stored in memory with state of the art computers, even taking account the possibilities of more powerful machines implementing Turing computation in a foreseeable future. This problem has been tackled in several ways.

The most obvious, most widely used way is *floating point computations*. The numbers are encoded by approximations which lie in a (non translation invariant) set of, so called, representable numbers. When any operator is applied, be it an arithmetical operation or a function (say an analytical function) of a number, the result, which is most often not representable, is approximated, in a way which is not always well specified in norms and may be implementation-dependent. During the execution of a complex numerical method, accumulated rounding-off errors can create convergence and instability problems for the method, even if the mathematical method is proved convergent, and stable with respect to initial values. But most of all, it is generally very difficult to prove that the result is correct with such methods.

Interval Analysis (see [AM00] for instance) consists in performing numerical computation by applying operators to interval, rather than to (encoding of) numbers. Each operation on an interval consists of operations, or possibly bounds calculation which are representable and

¹This work is supported by Agence Nationale pour la Recherche (A.N.R.), KIDICO (ANR 2010 BLANC 0205 02) project

calculable, for the boundaries of the interval. In this way, the technique allows to obtain bounds for solutions of equations or numerical problems. In many cases, however, the estimations are too loose to be really useful.

Exact Real arithmetic (see an introduction in [Vui87]) rests on the notion of a *computable number*, which consists in restricting constructions of real numbers by completion of the set \mathbb{Q} of rational numbers (by Cauchy Sequences, or by decreasing sequences of intervals), narrowing the construction to calculable (through some Turing Complete language) sequences. We can define an arithmetic on the sets of numbers which, under some restrictions related to computability, implement exactly numerical computation on calculable real numbers.

Symbolic Calculus consists in formal manipulations algebraic expressions, by rewriting formally the expressions according to the rules of the algebraic structures involved. This approach to exact real computation is generally considered a natural complement to exact real arithmetic, which can fill in gaps leaved by non-computable operations. This approach, however, is quite remote from numerical calculus as we purpose to do.

In *our approach*, we would define (hopefully) any numerical problem in a discrete form and in a continuous form. The discrete form of the problems would involve only integers (with some constant denominator or scale factor). Each of the considered operator would possibly change the rational scale factor (which can be stored), but otherwise be computable using only integers. Then, a *multigrid convergence* result would allow us to show that the integer-only computable solution of the discrete problem converges to the real valued solution of the continuous (and original) version of the problem. This avoids the accumulation of rounding-off errors typical of floating point computations.

To make such an approach succesfull, we would need to redefine so called digital, or discrete versions of all the objects involved in real analysis. Having made a short wish list (quadratic optimisation, gradient method, *PDE* integration, and so on), we found that the notion of a derivative is at the root of all these notions. *In this paper*, we propose a general definition of a digital derivative, and prove some multigrid convergence results.

In the framework of image and signal processing, as well as shape analysis, a common problem is to estimate derivatives of functions, or tangents and curvatures of curves and surfaces, when only some (possibly noisy) sampling of the function or curve is available from aquisition. This problem has been investigated through finite difference methods, scale-space ([KSW05], [Lin90]), and digital geometry ([CLL14], [CLL13], [KL09], [LVdV05], [LVdV07], [LCL14], [MB14], [PG11], [dVL09]).

In our previous work ([EMC11], [EM14], [MBF08], [FM09], [GMES13]) an approach to derivative estimation from digitized data was introduced. As in scale-space approaches, this approach is based on simple computations of *convolutions*. However, unlike scale-space methods, this approach is oriented towards integer-only models and algorithms, and is based on a discrete approach to analysis on \mathbb{Z} . Implementation of the convolution-based approach is straightforward. As far as the speed of convergence is concerned, the method for tangents of [MBF08] was proved in [EMC11] uniform worst-case $O(h^{2/3})$ for C^3 functions, where h is the size of the pixel. Moreover, our estimator allowed some (uniformly bounded or stochastic) noise. Furthermore, the method allowed to have a convergent estimation of higher order derivatives, and in particular a uniform $O(h^{4/9})$ estimation for the curvature of a generic curve.

The weakness of these results was the high complexity of computation. In this paper, we prove that the same convergence rates are allowed with a very low complexity by using sparse masks. This idea was previously used in connection with Taylor optimal masks ([EM14]). It

turns out that it is more efficient with sparse Binomial masks.

1.2 Outline

The paper is organized as follows. In Section 2, we present the notion of a digitization of a real-valued function, in which a discrete sequence of sample values approximates a real function. We also present a few noise models, including basic quantization, uniform noise or bias (depending on the scale), and stochastic noise.

In Section 3, we define digital derivatives. This notion is based on digital convolutions with masks which satisfy some identities involving sums (so-called moments). We also define some important characteristics, such as the convergence order of the mask. We also show some good properties, notably related to composition of digital derivatives by convolving the corresponding masks.

In Section 4, we begin to investigate errors for the purpose of proving multigrid convergence results. We distinguish the *sampling error*, which is due to the fact that we know only the values of the functions at a (locally) finite set of samples, from the *values error*, which is due to inaccurate input values (due to quantization, noise or bias). We also give upper bounds for both types of errors, for all the considered types of noise models.

In Section 5, we study some smoothing masks, i.e. some 0-derivative masks. These masks are important because, basically, any 1-derivative mask can be seen as a smoothing mask composed with a usual simple finite difference mask. The intuitive idea is that if we use the simple finite difference mask on noisy data, the importance of noise will be very detrimental to the results. By smoothing, which is nothing but a weighed average of values of the function, some of the noise cancels out (at least for a stochastic noise with expected value 0). We present three families of masks which we consider important, among the many masks on which we worked until now ([EMC11], [EM14], [MBF08], and [GMES13] in which we began to seek low computational complexity for convolutions). These three families of masks are the Binomial, Taylor-Optimal, and B -spline masks.

In Section 6, we establish multigrid convergence results of the digital derivatives toward the corresponding continuous derivatives of the original function, providing in addition a worst case rate of convergence. That section contains the main convergence results of this paper, namely uniform convergence results in the uniform noise or bias model, and convergence results of the standard deviation of our estimation in the stochastic model. Both convergence results, which hold for derivatives of any order ω , have a rate $O(h^{\alpha-\omega\alpha/k_0})$, where $0 < \alpha \leq 1$ depends on the noise and $k_0 \geq 2$ depends on the digital derivative operator. Furthermore, these convergence results apply to a general family of masks called skipping masks, which are basically obtained from any ω -derivative mask, by using a small mask with few values (for better computational complexity), but using a sparse set of sample values at points of a regular sub-grid.

In Section 7, we define three families (also called the Binomial, Taylor-Optimal, and B -spline masks) of skipping 1-derivative masks, with as corollary (from the results of Section 6) some multigrid convergence results obtained directly simply by checking the definition of a digital 1-derivative.

In Section 8.2, we apply these operators and results to some additional particular cases, such as second derivative, or the estimation of tangents and curvatures of parametrized curves.

In Section 9, we present a novel interpretation of digital convolutions as the computation of some points on a smooth curve, the differentiability of which depends on the kernel. We study in details the relationship between binomial kernels for derivatives of any order and the

corresponding derivatives of Bézier curves. We do the same for so-called B -spline kernels and B -spline curves. Since the B -spline case is somewhat simpler, we present it first. The B -spline kernels, which we introduce in this paper for the first time, seem a bit ad-hoc. We believe, however, that the B -spline case, just as the Bézier case, will be of great theoretical interest in the future.

The last section is devoted to experiments, the main purpose of which are manly to illustrate the difference between masks, but also that only a few (typically a small constant such as 10 to 20 points) values are needed to obtain a good estimation. An appendix also proposes different criteria for choosing a mask, depending on the considered application.

Contents

1	Introduction	1
1.1	Aim of the work	1
1.2	Outline	3
2	Digitization, Quantization, Noise Models	5
3	Digital Derivatives for Functions	6
4	Basic Error Decomposition and Upper Bounds	8
4.1	Errors Related to Sampling and to Input Values	8
4.2	Upper Bound for the Sampling Error	8
4.3	Upper Bound for the Values Error	9
5	Smoothing masks	10
5.1	Binomial Smoothing Masks	10
5.2	Taylor-Optimal Smoothing Masks	11
5.3	B -splines Smoothing Masks	12
6	Skipping Masks: Cheap Multigrid Convergence	13
6.1	Uniform Multigrid Convergence with Uniform Noise or Bias	13
6.2	Stochastic Multigrid Convergence with Stochastic Noise	14
7	Examples of 1-derivative masks	14
7.1	Binomial 1-Derivative Mask	14
7.2	Taylor-Optimal 1-Derivative Mask	15
7.3	B -spline 1-Derivative Mask	15
8	The cases of 2-Derivatives and Parametric Curves	16
8.1	Exemples of 2-derivative masks	16
8.2	Parametrized Curves	17
8.2.1	Binomial smoothing	17
8.2.2	Discrete tangent	17
8.2.3	Tangent Estimation	17
8.2.4	Discrete curvature	18
8.2.5	Curvature Estimation	18

9	Piecewise Polynomial Functions and Convolutions	19
9.1	Continuous B -Splines Revisited	19
9.1.1	Scaled B -Splines Curves Associated to a Sequence	19
9.1.2	Derivative of the B -splines Curves with Skipped Control Points	20
9.1.3	Relationship between scaled B -spline and convolutions	21
9.2	Bézier Curves with Skipped Control Points	21
9.2.1	Bézier Curves and Bernstein polynomials Basics	21
9.2.2	Scaled Bézier Function Associated to a Sequence	22
9.2.3	Derivative of the Scaled Bézier Function	23
9.2.4	Binomial Convolutions and Scaled Piecewise Bézier	24
10	Experiments	25
10.1	Smoothing	25
10.2	Derivatives	26
10.3	Curvature	27
11	Conclusion	28
A	Choosing a Mask	28
A.1	Criteria for Choosing a Mask	28
A.2	Moments of The Main Considered Masks	29
A.3	Smoothing Masks	29
A.4	First Order Derivative Masks	30
A.5	Second Order Derivatives	31

2 Digitization, Quantization, Noise Models

Functions for which domain and range are sub-ring of \mathbb{R} , without any assumption on the nature of these sets, are called *real functions*. We call *digital function* a function from \mathbb{Z} to \mathbb{Z} .

First we fix a model for the relationship between a continuous function and its digitization. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a real function and let $\Gamma : \mathbb{Z} \rightarrow \mathbb{Z}$ be a digital function. Let h be the digitization step (i.e. the size of a pixel). We introduce a (possibly noisy) digitization of f .

Definition 2.1 *The function Γ is a digitization of f with error ϵ_h with digitization step h on the interval I if for any integer i such that $ih \in I$ we have:*

$$h\Gamma(i) = f(ih) + \epsilon_h(i) \tag{1}$$

In order to apply this theory to signal analysis, we could distinguish two digitization steps for the domain and for the range, using the following definition: $h_2\Gamma(i) = f(ih_1) + \epsilon_{h_1, h_2}(i)$. The results concerning the estimation errors bounds may be easily generalized to this framework. However generalized convergence results may hardly be efficient because of the many parameters involved.

We consider the following particular models for the errors ϵ_h on the values:

- *Exact Values* In this model, the values are known exactly:

$$\epsilon_h = 0$$

Note that, although this model has been the most widely used in approximation theory (e.g. [Flo05], [Flo08]), this value error model is not very realistic from an Information Sciences point of view.

- *Uniform Noise (or Uniform Bias) on Values:* In this model, the error ϵ_h on the values is uniformly bounded by some constant which depends on the sampling step h . In our model, however, this bound can be asymptotically greater than h . Namely we assume here that

$$0 \leq |\epsilon_h(i)| \leq Kh^\alpha$$

where $0 < \alpha \leq 1$ and K is a positive constant. Note that this error can also have some bias, in the sense that the average noise value (or expected value) could be non-zero. This model has been considered by authors working on digital estimators based on thick digital lines.

- *Quantization of Values:* In this model, the errors ϵ_h on the values is uniformly bounded by $\frac{1}{2}h$. This is a particular case of uniform noise with $\alpha = 1$, and corresponds to the case when some basic quantization has been obtained by rounding-off the exactly known values of the function, for example for digital storage. This case is equivalent to $\Gamma(i) = \left\lfloor \frac{f(ih)}{h} \right\rfloor$. A variant is when quantization has been obtained by an integer part (floor case): $0 \leq \epsilon_h(i) < h$, which is equivalent to $\Gamma(i) = \lfloor \frac{f(ih)}{h} \rfloor$. This model is the most usual for digital estimators based on digital lines.
- *Stochastic Noise on Values:* In this model, the errors $\epsilon_h(i)$ on the different values are independent random variables with expected value 0 and standard deviation $\sigma(h)$, converging to 0 along with h . For convergence results, we shall suppose $\sigma(h) \leq Kh^\alpha$. This model, and generally some estimators using measure theory, have been used recently ([CLMT15]), but, to the best of our knowledge, without any multigrid convergence results.

3 Digital Derivatives for Functions

We introduce now a notion of digital derivatives.

Definition 3.1 (Digital Derivative) A digital ω -derivative mask is a sequence $\mathbf{u} = (u(i))_{i \in \mathbb{Z}} \in \mathbb{R}^{\mathbb{Z}}$ of finite support such that

$$\sum_{i \in \mathbb{Z}} i^k u(i) = 0 \text{ for } 0 \leq k < \omega \text{ and } \sum_{i \in \mathbb{Z}} i^\omega u(i) = (-1)^\omega \omega!$$

Its associated digital ω -derivative operator is the function $\Delta_{\mathbf{u}}$ with domain $\mathbb{R}^{\mathbb{Z}}$ and co-domain $\mathbb{R}^{\mathbb{Z}}$ defined by

$$\Delta_{\mathbf{u}}(\mathbf{v})(n) = (\mathbf{u} \star \mathbf{v})(n) = \sum_{i \in \mathbb{Z}} u(i)v(n-i).$$

Its convergence order is the greatest integer ρ such that for all integers k such that $1 + \omega \leq k \leq \rho$, we have $\sum_{i \in \mathbb{Z}} i^k u(i) = 0$.

We shall sometimes say a derivative mask as a shorthand for a digital derivative mask.

Remark 3.1 Note that if there is no integer k such that $1 + \omega \leq k$ and $\sum_{i \in \mathbb{Z}} i^k u(i) = 0$, then the convergence order ρ is equal to ω .

Remark 3.2 In the case when the values $u(i)$ are rational numbers and the sequence \mathbf{v} has integer values, then the values $\Delta_{\mathbf{u}}(\mathbf{v})(n)$ have rational values with denominator less than or equal to the GCD of the denominators of the $u(i)$.

Remark 3.3 Note that this definition of a derivative operator easily extends to some families of sequences with infinite support, under some convergence hypothesis (e.g. rapidly decreasing masks like a sampling of the Gaussian). So, this definition can be extended to include scale-space notions ([Lin90]).

For the sake of clarity, we shall denote by $\Delta_{\mathbf{u}}^{\omega}$ the associated digital derivation operator of \mathbf{u} when \mathbf{u} is a digital ω -derivative mask. Intuitively, and as appears in the proofs below, the order ρ of an ω -derivative operator is the number of coefficients of the Taylor development of the function $\frac{1}{h^{\omega-1}} \Delta_{\mathbf{u}}^{\omega}(f) - f^{(\omega)}$ which vanish, minus one. As shown below, the convergence order ρ will determine the speed of convergence of $\frac{1}{h^{\omega-1}} \Delta_{\mathbf{u}}^{\omega}(f)$ to $f^{(\omega)}$, which is to distinguish from the derivative order ω . Notice that the complexity of the computation is mainly dependent on the non zero coefficients of the mask, hence on the cardinality of its support (as long as it is computable). In [EMC11], this cardinal was an increasing function of h . One novelty of this work is considering masks with supports of constant cardinality, which is in practice so small as 10 or 12. Hence the computation needs no more than that number of products of integers.

We shall use the following generalization of the classical *finite difference operator* with order 1 and skipping step l :

Definition 3.2 The following operator Δ_l^+ is a digital 1-derivative operator:

$$\Delta_l^+(\mathbf{v})(n) = \frac{1}{l} (v(n+l) - v(n))$$

the mask of which is defined by $\Delta_l^+(-l) = \frac{1}{l}$, and $\Delta_l^+(0) = -\frac{1}{l}$ and $\Delta_l^+(i) = 0$ for others values of i .

Similarly, we denote by Δ_l^- the following derivative operator:

$$\Delta_l^-(\mathbf{v})(n) = \frac{1}{l} (v(n) - v(n-l))$$

As we could expect, the composition of two digital derivatives is also a digital derivative:

Proposition 3.1 Let $\mathbf{u} = (u(i))_{i \in \mathbb{Z}}$ be an ω -derivative mask and $\mathbf{v} = (v(i))_{i \in \mathbb{Z}}$ be an ω' -derivative mask. Then $\mathbf{u} \star \mathbf{v}$ is a $\omega + \omega'$ -derivative mask.

Proof. Let $k \leq \omega + \omega'$. Then

$$\begin{aligned} \sum_{n \in \mathbb{Z}} n^k (\mathbf{u} \star \mathbf{v})(n) &= \sum_{n \in \mathbb{Z}} (i + (n-i))^k \sum_{i \in \mathbb{Z}} u(i) v(n-i) \\ &= \sum_{n \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} \sum_{p=0}^k \binom{k}{p} i^p (n-i)^{k-p} u(i) v(n-i) \\ &= \sum_{p=0}^k \binom{k}{p} \left(\sum_{n \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} i^p u(i) (n-i)^{k-p} v(n-i) \right) \\ &= \sum_{p=0}^k \binom{k}{p} \sum_{j \in \mathbb{Z}} \sum_{i \in \mathbb{Z}} i^p u(i) j^{k-p} v(j) \\ &= \sum_{p=0}^k \binom{k}{p} \left(\sum_{j \in \mathbb{Z}} j^{k-p} v(j) \right) \left(\sum_{i \in \mathbb{Z}} i^p u(i) \right) \end{aligned}$$

This is zero except if $k = \omega + \omega'$ and in this case all the terms are zero except if $p = \omega$ and in this case the sum is $\binom{\omega+\omega'}{\omega'}(-1)^\omega \omega!(-1)^{\omega'} \omega'! = (-1)^{\omega+\omega'}(\omega + \omega')!$ \square

The following remark can be easily checked by induction on the degree of a polynomial:

Remark 3.4 *Our definition of a digital ω -derivative corresponds partially to the usual derivative when applied to polynomial sequences: let \mathbf{u} be a digital ω -derivative mask with convergence order ρ and let $v(i)$ be defined by i^k for a given non negative integer k ; then for $k \leq \rho$, we have*

$$(\Delta_{\mathbf{u}}^\omega((i^k)_{i \in \mathbb{Z}}))(n) = k(k-1)\dots(k-\omega+1)i^{k-\omega}.$$

4 Basic Error Decomposition and Upper Bounds

4.1 Errors Related to Sampling and to Input Values

In order to show that the digital ω -derivative of a digitization Γ of a real function f provides an estimate for the continuous derivative $f^{(\omega)}$ of f , we would like to evaluate, at each sample point, the difference between the digital derivative $\frac{1}{h^{\omega-1}}(\Delta_{\mathbf{u}}^\omega \star \Gamma)(n)$ of the digitized signal and the value of the usual ω -derivative $f^{(\omega)}(nh)$ of f . This difference may easily be decomposed from Equation (1) and Definition 3.1 as the sum

$$\frac{1}{h^{\omega-1}}(\Delta_{\mathbf{u}}^\omega \star \Gamma)(n) - f^{(\omega)}(nh) = ES_\omega(f, h, \Gamma, \mathbf{u}, n) + EV_\omega(f, h, \Gamma, \mathbf{u}, n) \quad (2)$$

where

$$ES_\omega(f, h, \Gamma, \mathbf{u}, n) = \left(\frac{1}{h^\omega} \sum_{i \in \mathbb{Z}} u(i) f((n-i)h) \right) - f^{(\omega)}(nh) \quad (3)$$

is called the *sampling error*, and

$$EV_\omega(f, h, \Gamma, \mathbf{u}, n) = \frac{1}{h^\omega} \sum_{i \in \mathbb{Z}} u(i) \epsilon_h(n-i) \quad (4)$$

is called the *values error*. As their names imply, the sampling error is due to the fact that we only know about the values of f at some grid points, and the values error is due to the fact that we do not know the exact values of f at sample points.

The sampling error is a real values sequence. Under the uniform bias hypothesis, the values error is also a real valued sequence, but under the stochastic hypothesis, the values error is a sequence of random variable.

4.2 Upper Bound for the Sampling Error

In the following lemma, we show that the sampling error can be bounded independently from the error on input values, using the mask values, the norm of the k_0 -derivative of f , and a power of the digitization step (namely $h^{k_0-\omega}$), where k_0 is (at least in the case when f is sufficiently regular) the convergence order of the mask \mathbf{u} . The consequence is some convergence results in the case when exact values of the function at sample points are known.

Lemma 4.1 Let f be C^k on \mathbb{R} , with $k \geq 1 + \omega$. Let \mathbf{u} be a digital ω -derivative mask with convergence order ρ . Let $k_0 = \inf\{k, 1 + \rho\}$. Let Γ be a digitization of f with step h on \mathbb{R} . Suppose that $f^{(k_0)}$ is bounded on \mathbb{R} . Then for all n ,

$$|ES_\omega(f, h, \Gamma, \mathbf{u}, n)| \leq h^{k_0 - \omega} \frac{\|f^{(k_0)}\|_\infty}{k_0!} \left(\sum_{i \in \mathbb{Z}} |i^{k_0} u(i)| \right)$$

Proof.

From Taylor formula, we first give an upper bound for the sampling error:

$$ES_\omega(f, h, \Gamma, \mathbf{u}, n) = \frac{1}{h^\omega} \sum_{i \in \mathbb{Z}} u(i) \left(\sum_{j=0}^{i=k_0-1} \frac{(-ih)^j}{j!} f^{(j)}(nh) + \frac{(-ih)^{k_0}}{k_0!} f^{(k_0)}(\zeta(i)) \right) - f^{(\omega)}(nh)$$

where $\zeta(i)$ lies between nh and $(n-i)h$. From the definition of the order, we get:

$$ES_\omega(f, h, \Gamma, \mathbf{u}, n) = h^{k_0 - \omega} \frac{(-1)^{k_0}}{k_0!} \left(\sum_{i \in \mathbb{Z}} i^{k_0} u(i) f^{(k_0)}(\zeta(i)) \right)$$

$$|ES_\omega(f, h, \Gamma, \mathbf{u}, n)| \leq h^{k_0 - \omega} \frac{\|f^{(k_0)}\|_\infty}{k_0!} \left(\sum_{i \in \mathbb{Z}} |i^{k_0} u(i)| \right) \quad \square$$

Example 4.1 From Definition 3.1, if f is C^2 and a 1-derivative mask \mathbf{u} has convergence order 1, then the sampling error is $O(h)$.

This is the case for the usual finite difference $(\Delta(f))(n) = f(n+1) - f(n)$.

Example 4.2 Suppose that f is C^3 and that a 1-derivative mask \mathbf{u} has convergence order 2, then the sampling error is $O(h^2)$.

This is the case for the usual symmetric finite difference $(\Delta(f))(n) = \frac{1}{2}(f(n+1) - f(n-1))$.

For a 2-derivative mask \mathbf{v} with convergence order 2, the sampling error for the second derivative is $O(h)$.

4.3 Upper Bound for the Values Error

The following lemma gives an upper bound for the error related to uniform noise or uniform bias on the values at sample points (see Section 2).

Lemma 4.2 Let f be a real function. Let \mathbf{u} be a digital ω -derivative mask with convergence order ρ . Let Γ be a digitization of f with step h on \mathbb{R} and uniform bias ϵ_h such that $\|\epsilon_h\|_\infty \leq Kh^\alpha$ with $0 < \alpha \leq 1$. Then for all n ,

$$|EV_\omega(f, h, \Gamma, \mathbf{u}, n)| \leq \frac{K}{h^{\omega - \alpha}} \left(\sum_{i \in \mathbb{Z}} |u(i)| \right)$$

Proof.

We straightforwardly give an upper bound for the values error:

$$|EV_\omega(f, h, \Gamma, \mathbf{u}, n)| \leq \frac{\|\epsilon_h\|_\infty}{h^\omega} \sum_{i \in \mathbb{Z}} |u(i)| \leq \frac{K}{h^{\omega - \alpha}} \left(\sum_{i \in \mathbb{Z}} |u(i)| \right).$$

\square

The following lemma gives an upper bound for the error related to statistic noise with expected values 0 on the values at sample points.

Lemma 4.3 *Let f be derivable on \mathbb{R} . Let \mathbf{u} be a digital ω -derivative mask. Let Γ be a digitization of f with step h on \mathbb{R} and stochastic noise on values ϵ_h with expected value 0 and standard deviation $\sigma(h)$. Then for all n , the random variable $\frac{1}{h^{\omega-1}}(\Delta_{\mathbf{u}}^{\omega} \star \Gamma)(n) - f^{(\omega)}(nh)$ has expected value $ES_{\omega}(f, h, \Gamma, \mathbf{u}, n)$ and standard deviation $\frac{\sigma(h)}{h^{\omega}} \sqrt{\sum_{i \in \mathbb{Z}} (u(i))^2}$. In other words, and roughly speaking, the global error is in this case statistically close to the sampling error.*

Proof. $\frac{1}{h^{\omega-1}}(\Delta_{\mathbf{u}}^{\omega} \star \Gamma)(n) - f^{(\omega)}(nh)$ is equal to the sum of the constant value $ES_{\omega}(f, h, \Gamma, \mathbf{u}, n)$ and the random variable defined by $EV_{\omega}(f, h, \Gamma, \mathbf{u}, n) = \frac{1}{h^{\omega}} \sum_{i \in \mathbb{Z}} u_i \epsilon_h(n - i)$. From linearity, its expected value is zero.

Since the random variables $\epsilon_h(n - i)$ are supposed to be independent and the support of \mathbf{u} is supposed to be finite, the square of the standard deviation of $EV_{\omega}(f, h, \Gamma, \mathbf{u}, n)$ is equal to the sum of the squares of the standard deviations of $\frac{u_i}{h^{\omega}} \epsilon_h(n - i)$ which are $\frac{|u_i|}{h^{\omega}} \sigma(h)$. \square

Note that for a fixed mask, the values error (or its standard deviation) generally does not converge to zero when h converges to 0. We shall propose below a way to make it tend to zero by adapting the mask to the digitization step (see Theorem 6.1 and Theorem 6.2 below).

5 Smoothing masks

A *smoothing mask* is a 0-derivative mask. In this case, the values $(\Delta_{\mathbf{u}}^0 \star \Gamma)(n)$ may be considered as partially de-noised values of $f(nh)$. More precisely, under the uniform hypothesis, according to our upper bounds from Lemmas 4.2 and 4.1, the difference between $h \cdot (\Delta_{\mathbf{u}}^0 \star \Gamma)(n)$ and $f(nh)$ cannot be considered not better (in order of magnitude) than the difference between $h \cdot \Gamma(n)$ and $f(nh)$. Moreover, Lemma 4.3 with $\omega = 0$ shows that, in the case of a stochastic noise on values, $h \cdot (\Delta_{\mathbf{u}}^0 \star \Gamma)(n) - f(nh)$ is a random variable, the expected value of which converges to 0 along with h , and its standard deviation is less or equal than the one of $h \cdot \Gamma(n) - f(nh)$.

Various choices for the smoothing masks may be considered. Here we focus on three families of masks.

5.1 Binomial Smoothing Masks

Definition 5.1 (Binomial Smoothing Masks) *Let m be a positive integer and i be an integer. The binomial smoothing mask with diameter $(2m - 1)$ and cardinality $2m$ is the finite sequence of dyadic rational numbers $\mathbf{B}_{2m-1}^{(0)}$ defined by $\mathbf{B}_{2m-1}^{(0)}(i) = \frac{1}{2^{2m-1}} \binom{2m-1}{m-i}$ if $-(m-1) \leq i \leq m$, and by $\mathbf{B}_{2m-1}^{(0)}(i) = 0$ in any other cases.*

The binomial smoothing mask with diameter $2m$ and cardinality $2m+1$ is the finite sequence of dyadic rational numbers $\mathbf{B}_{2m}^{(0)}$ defined by $\mathbf{B}_{2m}^{(0)}(i) = \frac{1}{2^{2m}} \binom{2m}{m-i}$ if $-m \leq i \leq m$, and by $\mathbf{B}_{2m}^{(0)}(i) = 0$ in any other cases.

A unified definition may be $\mathbf{B}_r^{(0)}(i) = \frac{1}{2^r} \binom{r}{\lfloor \frac{r}{2} \rfloor - i}$ if $-\lfloor \frac{r}{2} \rfloor \leq i \leq \lfloor \frac{r}{2} \rfloor$, and by $\mathbf{B}_r^{(0)}(i) = 0$ in all other cases.

Proposition 5.1 *The masks $\mathbf{B}_r^{(0)}$ for $r \geq 0$, are discrete smoothing masks with convergence order 0 if r is odd, and with convergence order 1 if r is even.*

Proof. We shall use the two following classical identities:

$$\sum_{i=0}^{i=n} i \binom{n}{i} = n2^{n-1} \quad \text{and} \quad \sum_{i=0}^{i=n} i^2 \binom{n}{i} = n(n+1)2^{n-2}.$$

$$\sum_{i \in \mathbb{Z}} \mathbf{B}_{2m-1}^{(0)}(i) = \sum_{i \in \mathbb{Z}} \mathbf{B}_{2m}^{(0)}(i) = 1 \text{ are well known.}$$

$$\sum_{i \in \mathbb{Z}} i \mathbf{B}_{2m}^{(0)}(i) = 0 \text{ come from parity.}$$

$$\sum_{i \in \mathbb{Z}} i \mathbf{B}_{2m-1}^{(0)}(i) = \frac{1}{2^{2m-1}} \sum_{i=-m+1}^m i \binom{2m-1}{m-i} = \frac{1}{2^{2m-1}} \sum_{j=0}^m (m-j) \binom{2m-1}{j}$$

$$\sum_{i \in \mathbb{Z}} i \mathbf{B}_{2m-1}^{(0)}(i) = m - \frac{1}{2}(2m-1) = \frac{1}{2}$$

$$\sum_{i \in \mathbb{Z}} i^2 \mathbf{B}_{2m}^{(0)}(i) = \frac{1}{2^{2m}} \sum_{i=-m}^m i^2 \binom{2m}{m-i} = \frac{1}{2^{2m}} \sum_{j=0}^m (m-j)^2 \binom{2m}{j}$$

$$\sum_{i \in \mathbb{Z}} i^2 \mathbf{B}_{2m}^{(0)}(i) = m^2 - 2m^2 + \frac{2m(2m+1)}{4} = \frac{m}{2} \quad \square$$

5.2 Taylor-Optimal Smoothing Masks

The following so called Taylor-optimal smoothing mask is chosen having the best order as possible for a given diameter:

Definition 5.2 (Taylor-Optimal Smoothing Masks) *Let m be a positive integer and i be an integer. The Taylor-optimal smoothing mask with diameter $2m$ and cardinality $2m$ is the finite sequence of rational numbers $\mathbf{O}_{2m}^{(0)}$ defined by $\mathbf{O}_{2m}^{(0)}(i) = (-1)^{i+1} \frac{\binom{m}{|i|}}{\binom{m+|i|}{m}} = (-1)^{i+1} \frac{\binom{2m}{m+|i|}}{\binom{2m}{m}}$ if $-m \leq i \leq m$, and by $\mathbf{O}_{2m}^{(0)}(i) = 0$ in any other cases.*

Proposition 5.2 *The mask $\mathbf{O}_{2m}^{(0)}$ is a digital smoothing mask with convergence order $2m-1$.*

Proof. Let $k \geq 0$. Let f be a polynomial and define the classical finite difference operator δ_+ by $\delta_+(f)(x) = f(x+1) - f(x)$. Iterating this operator leads to

$$\delta_+^{2m}(f)(x) = \sum_{i=0}^{2m} \binom{2m}{i} (-1)^{2m-i} f(x+i)$$

Applying this equality to the polynomials defined by $(x-n)^k$ for $0 \leq k \leq 2m-1$, we get

$$\sum_{i=0}^{2m} \binom{2m}{i} (-1)^i (i+x-m)^k = 0$$

Let now $x=0$, then for all $0 \leq k \leq 2m-1$ we have

$$0 = \sum_{i=0}^{2m} \binom{2m}{i} (-1)^i (i-m)^k = (-1)^m \sum_{j=-m}^{j=m} \binom{2m}{m+j} (-1)^j j^k$$

Hence for $k = 1$, we have $\sum_{j=-m, j \neq 0}^{j=m} (-1)^j \binom{2m}{m+j} = -\binom{2m}{m}$ and for $2 \leq k \leq 2m - 1$, we have $\sum_{j=-m, j \neq 0}^{j=m} (-1)^j j^k \binom{2m}{m+j} = 0$.

The same computation for the polynomial defined by $(x - m)^{2m}$ leads to $\sum_{j=-m, j \neq 0}^{j=m} (-1)^j j^{2m} \binom{2m}{m+j} = (-1)^m \neq 0$. \square

5.3 B -splines Smoothing Masks

Now, we define the B -spline Smoothing Kernels.

Definition 5.3 Let $s \in \mathbb{Z}$. The Cox-De-Boor functions $N_{s,r}$ (with knots $t_s = s$), for $r \in \mathbb{N}$, are defined inductively as follows.

$$N_{s,0}(t) = \begin{cases} 1 & \text{if } s \leq t < s + 1 \\ 0 & \text{otherwise.} \end{cases}$$

For $r \geq 1$, we use the induction formula (called the Cox-de Boor formula):

$$N_{s,r}(t) = \frac{t - s}{r} N_{s,r-1}(t) + \frac{s + r + 1 - t}{r} N_{s+1,r-1}(t)$$

Lemma 5.1 We have $N_{i,r}(t) = 0$ for $t \notin [i, i + r + 1[$. Besides, for all $t \in \mathbb{R}$ and $r \in \mathbb{N}$, we have the following properties:

Partition of Unity Property: $\sum_{i \in \mathbb{Z}} N_{i,r}(t) = 1$

Shift Property $N_{s,r}(t + 1) = N_{s-1,r}(t)$

Symmetry Property $N_{s,r}(s + \frac{r+1}{2} + t) = N_{0,r}(s + \frac{r+1}{2} - t)$

Proof. These properties are well known and follow immediately by induction from the Cox-de Boor formula. \square

Definition 5.4 (B -spline Smoothing Masks) Let us consider a positive integer $r \geq 0$. We introduce a convolution mask $(\mathbf{S}_r^{(0)}(i))_{i \in \mathbb{Z}}$ with support included in $\{-\lfloor \frac{r+1}{2} \rfloor, \dots, \lfloor \frac{r+1}{2} \rfloor\}$, based on the Cox-de Boor functions with degree r on \mathbb{R} , with (bi-infinite) nodal vector $\tau = (i)_{i \in \mathbb{Z}}$. The mask values are defined by $\mathbf{S}_r^{(0)}(i) = N_{0,r}(\frac{r+1}{2} + i) = N_{-i,r}(\frac{r+1}{2})$ if $-\lfloor \frac{r+1}{2} \rfloor \leq i \leq \lfloor \frac{r+1}{2} \rfloor$ and $\mathbf{S}_r^{(0)}(i) = 0$ otherwise. This mask is called the uniform B -spline smoothing mask with degree r .

Since $N_{0,r}(\frac{r+1}{2}) = 0$, the cardinality of the mask $\mathbf{S}_r^{(0)}$ is $2\lfloor \frac{r+1}{2} \rfloor$. The mask is always symmetric.

Proposition 5.3 The mask $\mathbf{S}_r^{(0)}$ is a 0-derivative mask (i.e. a smoothing mask) with convergence order 1.

Proof. The fact that $\mathbf{S}_r^{(0)}$ is a 0-derivative mask follows immediately from the partition of Unity property in Lemma 5.1 and Definition 3.1. The convergence order (same definition) equal to 1 follows from:

$$\sum_{i \in \mathbb{Z}} i \cdot \mathbf{S}_r^{(0)}(i) = \sum_{i \in \mathbb{Z}} i \cdot N_{0,r}\left(\frac{r+1}{2} + i\right)$$

The last sum is equal to zero from the symmetry property in Lemma 5.1. \square

The reader can find in Appendix A.3 some practical estimation for the constant involves in the asymptotic order sampling error for the three families of smoothing masks presented above.

6 Skipping Masks: Cheap Multigrid Convergence

The idea is adapting the mask to the step of digitization, in order to get $\frac{1}{h^{\omega-\alpha}} \left(\sum_{i \in \mathbb{Z}} |u(i)| \right)$ converging to zero along with h . For limiting the complexity of computation, we set the number of non zero coefficients of the mask fixed. This idea was independently used in [MB14] for length estimation. Other attempts have been considered in [GMES13].

Definition 6.1 (Skipping Masks) *Let \mathbf{u} be an ω -derivative mask with convergence order ρ , and $l \in \mathbb{N}^*$ be a skipping step. The corresponding ω -derivative l -skipping mask \mathbf{u}_l is defined by $\mathbf{u}_l(i) = \frac{1}{l^\omega} u(\frac{i}{l})$ if l divides i and 0 in any other cases. l is the length of the skips.*

As $\sum_{i \in \mathbb{Z}} i^k \mathbf{u}_l(i) = l^{k-\omega} \sum_{j \in \mathbb{Z}} j^k \mathbf{u}(j)$, the mask \mathbf{u}_l is an ω -derivative mask with same convergence order than \mathbf{u} . But as $\sum_{i \in \mathbb{Z}} |\mathbf{u}_l(i)| = \frac{1}{l^\omega} \sum_{j \in \mathbb{Z}} |\mathbf{u}(j)|$, it allows a convenient choice of l depending on h in order to get a zero converging values error.

6.1 Uniform Multigrid Convergence with Uniform Noise or Bias

Theorem 6.1 *Let \mathbf{u} be an ω -derivative mask with order ρ and \mathbf{u}_l the corresponding ω -derivative l -skipping mask with skips of length l . Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a C^k function and $f^{(k)}$ is bounded, $\alpha \in]0, 1]$, $K \in \mathbb{R}_+^*$ and $h \in \mathbb{R}_+^*$. Suppose $\Gamma : \mathbb{Z} \rightarrow \mathbb{Z}$ is such that $|h\Gamma(i) - f(hi)| \leq Kh^\alpha$ for all i . Let $k_0 = \text{Min}\{k, 1 + \rho\}$. Then for $k \geq 1 + \omega$ and $l(h) = \lfloor h^{-1+\alpha/k_0} \rfloor$,*

$$\text{we have } \left| \left(\frac{1}{h^{\omega-1}} \Delta_{\mathbf{u}_{l(h)}} \star \Gamma \right) (n) - f^{(\omega)}(nh) \right| \in O(h^{\alpha-\omega\alpha/k_0}),$$

$$\text{and, for } k = \omega \text{ and } l \text{ such that } \lim_{h \rightarrow 0} h.l(h) = 0 \text{ and } \lim_{h \rightarrow 0} \frac{h^\alpha}{((h.l(h))^\omega)} = 0,$$

$$\text{we have } \left| \left(\frac{1}{h^{\omega-1}} \Delta_{\mathbf{u}_{l(h)}} \star \Gamma \right) (n) - f^{(\omega)}(nh) \right| \in o(1).$$

Proof. First, we give an upper bound for the values error. From Lemma 4.2 and definitions, we have $|EV(f, h, \Gamma, \mathbf{u}_{l(h)}, n)| \leq \frac{K}{l(h)^\omega h^{\omega-\alpha}} \sum_{j \in \mathbb{Z}} |u(j)|$. If $l(h) = \lfloor h^{-1+\alpha/k_0} \rfloor$, it is easy to verify

$$\text{that } \frac{1}{l(h)^\omega h^{\omega-\alpha}} \leq \frac{h^{\alpha-\omega\alpha/k_0}}{1 - h^{\omega-\omega\alpha/k_0}} \text{ which is } O(h^{\alpha-\omega\alpha/k_0}).$$

We now turn to the sampling error:

- Let $k \geq 1 + \omega$. From Lemma 4.1 and definitions, the sampling error is less than $(l(h)h)^{k_0-\omega} \frac{\|f^{(k_0)}\|_\infty}{k_0!} \sum_{j \in \mathbb{Z}} |j^{k_0} u(j)|$. If $l(h) = \lfloor h^{-1+\alpha/k_0} \rfloor$, it is easy to verify that $|ES(f, h, \Gamma, \mathbf{u}_{l(h)}, n)| \leq h^{\alpha-\omega\alpha/k_0} \frac{\|f^{(k_0)}\|_\infty}{k_0!} \sum_{j \in \mathbb{Z}} |j^{k_0} u(j)|$.
- Let f be C^ω on \mathbb{R} . From Taylor formula, $ES(f, h, \Gamma, \mathbf{u}_{l(h)}, n)$ turn to be equal to $\sum_{j \in \mathbb{Z}} j^\omega u(j) \epsilon((n - jp(h))h)$, where ϵ is a function such that $\lim_{x \rightarrow 0} \epsilon(x) = 0$. \square

The multiplicative constants in the first part of Theorem 6.1 depends on various parameters. Some depend on informations about the data: a bound h_{max} of the digitization step, or $\|f^{(k_0)}\|_\infty$ or K depending on the model of the noise. Some others depend on the mask: $\sum_{i \in \mathbb{Z}} |i^k u_i|$ is named the k -moment of the mask. In the case when $k \geq 1 + \omega$, the error is bounded by

$$\left(\frac{K}{1 - h_{max}^{\alpha-\omega\alpha/k_0}} \sum_{i \in \mathbb{Z}} |u_j| + \frac{\|f^{(k_0)}\|_\infty}{k_0!} \sum_{i \in \mathbb{Z}} |j^{k_0} u_j| \right) h^{\alpha-\omega\alpha/k_0}$$

For practical purposes, it could be useful to choose the mask in order to minimize this constant. In Appendix A.2, we provide formulas for some of our masks, and table of values for others.

6.2 Stochastic Multigrid Convergence with Stochastic Noise

Theorem 6.2 *Let \mathbf{u} be a ω -derivative mask with convergence order ρ and \mathbf{u}_l the corresponding ω -derivative l -skipping mask. Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a C^k function, with $k \geq 1 + \omega$, and $f^{(k)}$ is bounded, $\alpha \in]0, 1]$, $K \in \mathbb{R}_+^*$ and $h \in \mathbb{R}_+^*$. Let Γ be a digitization of f with step h and a stochastic noise ϵ_h with expected value 0 and standard deviation $\sigma(h) \leq Kh^\alpha$. Then for $k_0 = \text{Min}\{k, 1 + \rho\}$ and $l(h) = \lfloor h^{1-\alpha/k_0} \rfloor$, the random variable $\left(\frac{1}{l(h)^\omega} \Delta_{\mathbf{u}_l(h)} \star \Gamma \right) (n) - f^{(\omega)}(nh)$ has an expected value and a standard deviation which are $O(h^{\alpha-\omega\alpha/k_0})$.*

The proof is similar to that of Theorem 6.1.

7 Examples of 1-derivative masks

It should be noted that all the masks defined in this section, which are digital 1-derivative ($\omega = 1$) skipping masks, can be proven to have a multigrid convergence property either from Theorem 6.1 or from Theorem 6.2 (depending on the model of errors on values). We remind the reader that a more precise idea of the speed of convergence for the different masks can be found in Appendix A.2.

7.1 Binomial 1-Derivative Mask

Definition 7.1 (Binomial skipping masks) *Let m and l be positive integers and i be an integer. The binomial 1-derivative l -skipping mask with diameter $2ml$ and cardinality $2m + 1$ for discrete first order derivation is the finite sequence of rational numbers $\mathbf{B}_{l,2m-1}^{(1)}$ defined by $\mathbf{B}_{l,2m-1}^{(1)}(i) = \frac{1}{l^{2m-1}} \left(\binom{2m-1}{m-1-\frac{i}{l}} - \binom{2m-1}{m-\frac{i}{l}} \right)$ if l divides i and $-ml \leq i \leq ml$, and by $\mathbf{B}_{l,2m-1}^{(1)}(i) = 0$ in any other cases.*

Remark 7.1 *Note that $\mathbf{B}_{l,2m-1}^{(1)}$ is the convolution product $\Delta_l^+ \star \mathbf{B}_{l,2m-1}^{(0)}$. From Proposition 3.1, this is a general method for building a derivative mask from a smoothing mask.*

Proposition 7.1 *The masks $\mathbf{B}_{l,2m-1}^{(1)}$ are discrete 1-derivative masks of convergence orders 2.*

Proof. We prove here the following equalities:

$$\begin{aligned} \sum_{i \in \mathbb{Z}} \mathbf{B}_{l,2m-1}^{(1)}(i) &= 0 \\ \sum_{i \in \mathbb{Z}} i \mathbf{B}_{l,2m-1}^{(1)}(i) &= -1 \\ \sum_{i \in \mathbb{Z}} i^2 \mathbf{B}_{l,2m-1}^{(1)}(i) &= 0 \\ \sum_{i \in \mathbb{Z}} i^3 \mathbf{B}_{l,2m-1}^{(1)}(i) &= \frac{1-3m}{2} l^2 \end{aligned}$$

Let us notice that for all k , we have:

$$\begin{aligned} \sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{l^{k-1}}{2^{2m-1}} \sum_{j=-m}^{j=m} j^k \left(\binom{2m-1}{m-1-j} - \binom{2m-1}{m-j} \right) \\ \text{and} \\ \sum_{j=-m}^{j=-1} j^k \left(\binom{2m-1}{m-1-j} - \binom{2m-1}{m-j} \right) &= \sum_{j=1}^{j=m} (-j)^k \left(\binom{2m-1}{m-1+j} - \binom{2m-1}{m+j} \right) \\ \sum_{j=-m}^{j=-1} j^k \left(\binom{2m-1}{m-1-j} - \binom{2m-1}{m-j} \right) &= \sum_{j=1}^{j=m} j^k \left(\binom{2m-1}{m-j} - \binom{2m-1}{m-1-j} \right) \end{aligned}$$

Hence, for the even values of k , we have $\sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)}(i) = 0$.

For the odd values of k , we have:

$$\begin{aligned}\sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{p^{k-1}}{2^{2m-1}} \left(\sum_{j=-m}^{j=m-1} j^k \binom{2m-1}{m-1-j} - \sum_{j=-m+1}^{j=m} j^k \binom{2m-1}{m-j} \right) \\ \sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{p^{k-1}}{2^{2m-1}} \left(\sum_{j=-m+1}^{j=m} (j-1)^k \binom{2m-1}{m-j} - \sum_{j=-m}^{j=m-1} j^k \binom{2m-1}{m-j} \right) \\ \sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{p^{k-1}}{2^{2m-1}} \sum_{j=-m}^{j=m} ((j-1)^k - j^k) \binom{2m-1}{m-j}\end{aligned}$$

For $k = 1$, we get $\sum_{i \in \mathbb{Z}} i \mathbf{B}_{l,2m-1}^{(1)} = -1$.

For $k = 3$, we use previous results:

$$\begin{aligned}\sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{p^{k-1}}{2^{2m-1}} \sum_{j=-m}^{j=m} (-3j^2 + 3j - 1) \binom{2m-1}{m-j} \\ \sum_{i \in \mathbb{Z}} i^k \mathbf{B}_{l,2m-1}^{(1)} &= \frac{p^{k-1}}{2^{2m-1}} (-3m \cdot 2^{2m-2} + 3 \cdot 2^{2m-2} - 2^{2m-2})\end{aligned}$$

□

Proposition 7.2 *The masks $\mathbf{B}_{l,2m}^{(1)}$ are discrete 1-derivative masks of convergence orders 1.*

The proof could be derived from Remark 7.1 and Proposition 3.1, or we could follow the first steps of the proof of Proposition 7.1.

7.2 Taylor-Optimal 1-Derivative Mask

Definition 7.2 (Taylor-Optimal l -skipping masks) *Let m and l be positive integers and i be an integer. The Taylor-optimal 1-derivative l -skipping mask with diameter $2ml$ and cardinality $2m + 1$ for discrete first order derivation is the finite sequence of rational numbers $\mathbf{O}_{l,m}^{(1)}$ defined by $\mathbf{O}_{l,m}^{(1)}(i) = \frac{(-1)^i}{i} \frac{\binom{m}{\lfloor \frac{i}{l} \rfloor}}{\binom{m+\lfloor \frac{i}{l} \rfloor}{m}} = \frac{(-1)^i}{i} \frac{\binom{2m}{m+\lfloor \frac{i}{l} \rfloor}}{\binom{2m}{m}}$ if l divides i and $-ml \leq i \leq ml$, and by $\mathbf{O}_{l,m}^{(1)}(i) = 0$ in any other cases.*

Let us note that the complexity of the computation depends mainly on the parameter m and a few on the parameter l .

Proposition 7.3 $\mathbf{O}_{l,m}^{(1)}$ are discrete 1-derivative mask with convergence order $2m$.

The proof follows easily from Proposition 5.2.

7.3 B-spline 1-Derivative Mask

Definition 7.3 (B-spline Derivative Kernel) *Let us consider two integers $r \geq 0$ and $l > 0$. We introduce a convolution mask $\mathbf{S}_{l,r}^{(1)}$, with support in $\{-\lfloor \frac{lr}{2} + 1 \rfloor, \dots, \lfloor \frac{lr}{2} + 1 \rfloor\}$, based on the uniform Cox-de Boor functions with degree $r - 1$ on \mathbb{R} . The mask values are defined by $\mathbf{S}_{l,r}^{(1)}(i) = \frac{1}{l} (N_{0,r-1}(\frac{r}{2} + \frac{i}{l}) - N_{0,r-1}(\frac{r}{2} + \frac{i}{l} - 1))$ if l divides i and $-\lfloor \frac{lr}{2} + 1 \rfloor \leq i \leq \lfloor \frac{lr}{2} + 1 \rfloor$ and $\mathbf{S}_{l,r}^{(1)}(i) = 0$ otherwise.*

This mask is called the B-spline uniform l -skipping 1-derivative mask with degree $r - 1$.

Remark 7.2 *Note that $\mathbf{S}_{l,r}^{(1)} = \Delta_l^- * \mathbf{S}_{l,r-1}^{(0)}$, where $\mathbf{S}_{l,r-1}^{(0)}$ denotes the l -skipping smoothing mask associated to the uniform B-spline smoothing mask $\mathbf{S}_{r-1}^{(0)}$ with degree $r - 1$.*

Proposition 7.4 *The mask $\mathbf{S}_{l,r}^{(1)}$ is a 1-derivative mask with convergence order 1.*

The proof could be derived from Remark 7.2 and Proposition 3.1, but we present a hand made proof.

Proof. The definition of a 1-derivative mask with convergence order 1 consists of two sum equalities. Hence the proof consists of two steps:

First Step.

$l \sum_{i \in \mathbb{Z}} \mathbf{S}_{l,r}^{(1)}(i) = \sum_{i \in \mathbb{Z}} N_{0,r-1} \left(\frac{r}{2} + \frac{i}{l} + 1 \right) - N_{0,r-1} \left(\frac{r}{2} + \frac{i}{l} - 1 \right) = 0$. The last equality follows from change of index $i \leftrightarrow (i - l)$ in one of the sums.

Second Step.

$$\begin{aligned} l \sum_{i \in \mathbb{Z}} i \cdot \mathbf{S}_{l,r}^{(1)}(i) &= \sum_{i \in \mathbb{Z}} i \cdot [N_{0,r-1} \left(\frac{r}{2} + \frac{i}{l} \right) - N_{0,r-1} \left(\frac{r}{2} + \frac{i}{l} - 1 \right)] \\ &= \sum_{i \in \mathbb{Z}} i \cdot N_{0,r-1} \left(\frac{r-1}{2} + \frac{i}{l} \right) - \sum_{i \in \mathbb{Z}} (i+l) \cdot N_{0,r-1} \left(\frac{r-1}{2} + \frac{i}{l} \right) \\ &= -l \sum_{i \in \mathbb{Z}} N_{0,r-1} \left(\frac{r-1}{2} + \frac{i}{l} \right) = -l \end{aligned}$$

□

The following methods for building derivative masks from a smoothing mask \mathbf{u} of order ρ are worth noting:

1. Let \mathbf{v} be the sequence defined by $v(0) = \frac{1}{\sum_{j \in \mathbb{Z}, j \neq 0} u(j)} \sum_{j \in \mathbb{Z}, j \neq 0} \frac{1}{j} u(j)$ and $v(i) = \frac{-u(i)}{i \sum_{j \in \mathbb{Z}, j \neq 0} u(j)}$ for $i \neq 0$. Then, \mathbf{v} is a 1-derivative mask with order $1 + \rho$ if and only if $u(0) = 0$.
2. From Proposition 3.1 and Definition 3.2, we can easily define and $(\omega + 1)$ -derivative mask from any ω -derivative mask by convolution with Δ_l^+ .

Evaluating the errors needs to evaluate $\sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{l,2m-1}^{(1)}(i)|$ and $\sum_{i \in \mathbb{Z}} |i^k \mathbf{O}_{l,m}^{(1)}(i)|$ (see the proof of Theorem 6.1). See Appendix A.4 for the most useful such estimations.

8 The cases of 2-Derivatives and Parametric Curves

8.1 Examples of 2-derivative masks

Concerning second order derivatives, we can consider similar l -skipping masks. It is worth noting the following general method: Let \mathbf{u} and \mathbf{v} be two 1-derivative masks, then $\mathbf{u} \star \mathbf{v}$ is a 2-derivative mask (see Proposition 3.1).

Definition 8.1 *Let m and l be positive integers and i an integer. The Taylor-optimal l -skipping mask with diameter $2ml$ and cardinality $2m + 1$ for discrete second order derivation is the finite sequence of rational numbers $\mathbf{O}_{l,m}^{(2)}$ defined by $\mathbf{O}_{l,m}^{(2)}(i) = \frac{-2}{li} \mathbf{O}_{l,m}^{(1)}(i)$ if l divides i and $-ml \leq i \leq ml$ and $i \neq 0$ and $\mathbf{O}_{l,m}^{(2)}(0)$ is such that $\sum_{i=-ml}^{ml} \mathbf{O}_{l,m}^{(2)}(i) = 0$, and 0 for all other values of i .*

Lemma 8.1 $\mathbf{O}_{l,m}^{(2)}$ is a l -skipping 2-derivative mask with convergence order $1 + 2m$.

The proof follows easily from Proposition 5.2.

Definition 8.2 *Let m be a positive integers and i be an integer.*

The binomial smoothing mask with diameter $(2m+2)l$ and cardinality $2m+3$ is the finite sequence of rational numbers $\left(\mathbf{B}_{l,2m-1}^{(2)} \right)$ defined by $\mathbf{B}_{l,2m-1}^{(2)}(i) = \frac{1}{l^2 2^{2m-1}} \left(\binom{2m-1}{m+1-\frac{i}{l}} - 2 \binom{2m-1}{m-\frac{i}{l}} + \binom{2m-1}{m-1-\frac{i}{l}} \right)$ if $-(m+1)l \leq i \leq (m+1)l$ and l divides i , and by $\mathbf{B}_{l,2m-1}^{(2)}(i) = 0$ in any other cases.

See Appendix A.5 for estimations of constants in the convergence order.

8.2 Parametrized Curves

Let $k \geq 1$. We assume that a planar closed C^k -parametrized curve C (i.e. the parametrization is periodic) is given together with a family of parametrized digital curves (Σ_h) with Σ_h contained in a tube with radius $H(h)$ around C . A parametrization of C is denoted by $\mathbf{g} = (g_1, g_2) : \mathbb{R} \mapsto \mathbb{R}^2$, where g_1 and g_2 are C^1 and periodic with a common period. The parametrization of Σ_h is denoted by $\mathbf{\Gamma}_h = (\Gamma_1, \Gamma_2) : \mathbb{Z} \mapsto \mathbb{Z}^2$.

8.2.1 Binomial smoothing

Let $\mathbf{P} = (P_0, \dots, P_n)$ be a $(n + 1)$ -tuple of control points. Let us denote a Bézier curve $B_{\mathbf{P}}(t) = \sum_{i=0}^n \binom{n}{i} t^i (1-t)^{n-i} P_i$, where $t \in [0, 1]$ (See the Section 9.2). Thus we have

$$B_{(\mathbf{\Gamma}(i_0-m+\zeta), \dots, \mathbf{\Gamma}(i_0+m))}(\frac{1}{2}) = \left(\left(\Delta_{\mathbf{B}_{2m-\zeta}^{(0)}}^0(\Gamma_1) \right) (i_0), \left(\Delta_{\mathbf{B}_{2m-\zeta}^{(0)}}^0(\Gamma_2) \right) (i_0) \right)$$

where $\zeta = 0$ or 1 .

8.2.2 Discrete tangent

Here we estimate the tangent at a point of C by a digital tangent at a not too far point of Σ_h . To this purpose, we define the digital derivative in a natural way with the digital derivative of the coordinates. Note that this definition needs no hypothesis concerning the parametrized curve. Moreover, we could theoretically use different masks for the two coordinates. For sake of simplicity, we shall use the same for both.

Definition 8.3 *Let \mathbf{u} be a 1-derivative mask. Let $\mathbf{\Gamma} = (\Gamma_1, \Gamma_2) : \mathbb{Z} \mapsto \mathbb{Z} \times \mathbb{Z}$ be a digital parametrization of a digital curve Σ .*

The digital derivative of $\mathbf{\Gamma}$ at point $\mathbf{\Gamma}(n)$ is

$$\left((\Delta_{\mathbf{u}}^1(\Gamma_1)) (n), (\Delta_{\mathbf{u}}^1(\Gamma_2)) (n) \right)$$

It is denoted by $(\Delta_{\mathbf{u}}^1(\mathbf{\Gamma})) (n)$. When this vector is nonzero, the corresponding discrete tangent at $\mathbf{\Gamma}(n)$ is the real line going through $\mathbf{\Gamma}(n)$ directed by $(\Delta_{\mathbf{u}}^1(\mathbf{\Gamma})) (n)$.

Let us notice that the asymmetric l -skipping binomial case corresponds to the derivative of the Bézier curve with control points $(\mathbf{\Gamma}(i_0 - (m - 1)l), \dots, \mathbf{\Gamma}(i_0 + ml))$ computed at the medial-point of parameter $\frac{1}{2}$.

8.2.3 Tangent Estimation

Theorem 8.1 *Let \mathbf{u} be a smoothing mask with convergence order ρ . Let $\mathbf{g} = (g_1, g_2)$ be a C^k parametrization of a closed curve Σ . Let $k_0 = \inf\{k, 1 + \rho\}$, $\alpha \in]0, 1]$, $K \in \mathbb{R}_+^*$ and $h \in \mathbb{R}_+^*$ with $h \leq \frac{1}{2}$. Let $\mathbf{\Gamma}_h = (\Gamma_{h,1}, \Gamma_{h,2}) : \mathbb{Z} \mapsto \mathbb{Z} \times \mathbb{Z}$ be a digital parametrization of Σ . Suppose that for all i we have $\|\mathbf{g}(ih) - h \cdot \mathbf{\Gamma}_h(i)\|_{\infty} \leq Kh^{\alpha}$. Suppose that $g_c^{(k_0)}$ is bounded for $c = 1, 2$.*

Then $\|\Delta_{\mathbf{u}}^1(\mathbf{\Gamma}_h)(n) - \mathbf{g}'(nh)\|_{\infty}$ is less than

$$h^{k_0-1} \frac{\text{Max} \left\{ \|g_1^{(k_0)}\|_{\infty}, \|g_2^{(k_0)}\|_{\infty} \right\}}{k_0!} \left(\sum_{i \in \mathbb{Z}} |i^{k_0} u(i)| \right) + \frac{K}{h^{1-\alpha}} \left(\sum_{i \in \mathbb{Z}} |u(i)| \right)$$

The proof follows directly from Lemma 4.2 and Lemma 4.1.

Theorem 8.2 *Let \mathbf{u} be a 1-derivative mask with convergence order ρ and \mathbf{u}_l the corresponding 1-derivative l -skipping mask. Under the hypothesis of Theorem 8.1, for $k \geq 1$ and $l(h) = \lfloor h^{-1+\alpha/k_0} \rfloor$ we have $\left\| \Delta_{\mathbf{u}_l(h)}^1(\Gamma_h)(n) - \mathbf{g}'(nh) \right\|_\infty \in O(h^{\alpha-\alpha/k_0})$, and for $k = 1$ and $l(h) = \lfloor h^{-1+\alpha/2} \rfloor$, we have $\left\| \Delta_{\mathbf{u}_l(h)}^1(\Gamma_h)(n) - \mathbf{g}'(nh) \right\|_\infty \in o(1)$*

The proof is straightforward.

The hypothesis of Theorem 8.1 is stronger than the digitization being simply contained in a tube, because not only the discrete curve must be close to the continuous curve, but the parametrization of the digitization must be close to the parametrization of the continuous curve. This was the reason for introducing Pixel-Length parametrizations in [EMC11].

8.2.4 Discrete curvature

We estimate the curvature at a point of C by a digital curvature at a not too far point of Σ_h . The definition of the digital curvature coincides with the classical formula for continuous curves.

Definition 8.4 *Let \mathbf{u} and \mathbf{v} be respectively a 1-derivative mask and 2-derivative mask. Let Σ be a digital parametrized curve of parametrization $\Gamma = (\Gamma_1, \Gamma_2)$. The corresponding discrete curvature at $\Gamma(n)$ is*

$$\frac{(\Delta_{\mathbf{v}}^2(\Gamma_1))(n) (\Delta_{\mathbf{u}}^1(\Gamma_2))(n) - (\Delta_{\mathbf{u}}^1(\Gamma_1))(n) (\Delta_{\mathbf{v}}^2(\Gamma_2))(n)}{\left(((\Delta_{\mathbf{u}}^1(\Gamma_1))(n))^2 + ((\Delta_{\mathbf{u}}^1(\Gamma_2))(n))^2 \right)^{3/2}}$$

when the vector $\Delta_{\mathbf{u}}^1(\Gamma)(n)$ is non-zero. It will be denoted by $C_{\mathbf{u},\mathbf{v}}(\Gamma)(n)$.

Notice that this definition needs no more hypothesis concerning the parametrized curve.

8.2.5 Curvature Estimation

We shall denote $\gamma_{\mathbf{g}}(t)$ the curvature of a curve parametrized by \mathbf{g} at the point of parameter t . The following convergence theorem required some additional hypothesis concerning the regularity of the considered point.

Theorem 8.3 *Let \mathbf{u} be a 1-derivative mask with convergence order $\rho(\mathbf{u})$ and \mathbf{u}_l the corresponding 1-derivative l -skipping mask.*

Let \mathbf{v} be a 2-derivative mask with convergence order $\rho(\mathbf{v})$ and \mathbf{v}_l the corresponding 2-derivative l -skipping mask.

Let $\mathbf{g} = (g_1, g_2)$ be a C^k parametrization of a simple closed curve, with $k \geq 2$.

Let $\Gamma = (\Gamma_1, \Gamma_2) : \mathbb{Z} \mapsto \mathbb{Z} \times \mathbb{Z}$ be a digital parametrization of a digital curve Σ .

Let $k_0 = \inf\{k, 1 + \rho(\mathbf{u}), 1 + \rho(\mathbf{v})\}$, $\alpha \in]0, 1]$, $K \in \mathbb{R}_+^$ and $h \in \mathbb{R}_+^*$ and $h \leq \frac{1}{2}$. Suppose that for all i we have $\|\mathbf{g}(nh) - h.\Gamma(n)\|_\infty \leq Kh^\alpha$. Suppose that $g_c^{(k_0)}$ is bounded for $c = 1, 2$. Suppose that for $t \in [t_1, t_2]$, the vector $\mathbf{g}'(t) \neq \mathbf{0}$.*

Then, there is some $h_0 > 0$ and two functions $n_1, n_2 : (0; +\infty) \mapsto \mathbb{Z}$ such that for $l(h) = \lfloor h^{-1+\alpha/k_0} \rfloor$, for $h \geq h_0$ and $n \in [n_1(h), n_2(h)] \cap \mathbb{Z}$, we have

$$\left\| C_{\mathbf{u}_l(h), \mathbf{v}_l(h)}(\Gamma)(n) - \gamma_{\mathbf{g}}(nh) \right\| \in O(h^{\alpha-2\alpha/k_0})$$

Proof. Let us introduce the function defined on $\mathbb{R}^2 - \{(0, 0)\} \times \mathbb{R}^2$ by

$$F(X, Y, Z, T) = \frac{ZY - XT}{(X^2 + Y^2)^{\frac{3}{2}}}$$

Hence $C_{\mathbf{u}_1(\mathbf{h}), \mathbf{v}_1(\mathbf{h})}(\mathbf{\Gamma})(n) = F(\Delta_{\mathbf{u}_1(\mathbf{h})}^1(\mathbf{\Gamma})(n), \Delta_{\mathbf{v}_1(\mathbf{h})}^2(\mathbf{\Gamma})(n))$ and $\gamma_{\mathbf{g}}(nh) = F(\mathbf{g}'(nh), \mathbf{g}''(nh))$.

The differential of F is easily seen to be bounded on any compact \mathbf{K} in $\mathbb{R}^2 - \{(0, 0)\} \times \mathbb{R}^2$. Let $M_{\mathbf{K}}$ be one of its upper bounds on \mathbf{K} . Then for \mathbf{a} and \mathbf{b} in \mathbf{K} , we have $\|F(\mathbf{a}) - F(\mathbf{b})\|_{\infty} \leq M_{\mathbf{K}}\|\mathbf{a} - \mathbf{b}\|_{\infty}$.

We claim that, under the hypothesis of the theorem, we know that there is a compact \mathbf{K}_1 in $\mathbb{R}^2 - \{(0, 0)\}$ such that for $t \in [t_1, t_2] \subset \mathbb{R}$, the derivative $\mathbf{g}'(t)$ lies in \mathbf{K}_1 . Indeed $\mathbf{K}_1 = \{\mathbf{x} \in \mathbb{R}^2; \alpha \leq \|\mathbf{x}\| \leq \beta\}$, where $\alpha = \text{Inf}\{\|\mathbf{g}'(t)\|; t \in [t_1; t_2]\}$ and $\beta = \text{Max}\{\|\mathbf{g}'(t)\|; t \in [t_1; t_2]\}$ is convenient.

Then from Theorem 8.2, we can find a compact \mathbf{K}_2 in $\mathbb{R}^2 - \{(0, 0)\}$ and a fixed h_1 such that for $h \geq h_1$ and $n \in [\lceil \frac{t_1}{h} \rceil, \lfloor \frac{t_2}{h} \rfloor] \cap \mathbb{Z}$ the derivatives $\mathbf{g}'(nh)$ and $(\Delta_{\mathbf{u}_1(\mathbf{h})}^1(\mathbf{\Gamma})(n))$ ly in \mathbf{K}_2 .

Moreover, under previous conditions, the second order derivatives $\mathbf{g}''(nh)$ and $(\Delta_{\mathbf{v}_1(\mathbf{h})}^2(\mathbf{\Gamma})(n))$ lie in a compact \mathbf{K}_3 of \mathbb{R}^2 .

From Theorem 6.1, there is a constant χ and a fixed h_2 such $h \geq h_2$, we have

$$\left\| (\Delta_{\mathbf{u}_1(\mathbf{h})}^1(\mathbf{\Gamma})(n), \Delta_{\mathbf{v}_1(\mathbf{h})}^2(\mathbf{\Gamma})(n)) - (\mathbf{g}'(nh), \mathbf{g}''(nh)) \right\|_{\infty} \leq \chi \cdot h^{\alpha - 2\alpha/k_0}$$

Hence for $h \geq \max\{h_1, h_2\}$ and $n \in [n_1(h), n_2(h)] \cap \mathbb{Z}$, we have

$$\left\| C_{\mathbf{u}_1(\mathbf{h}), \mathbf{v}_1(\mathbf{h})}(\mathbf{\Gamma})(n) - \gamma_{\mathbf{g}}(nh) \right\|_{\infty} \leq \chi \cdot M_{\mathbf{K}_2 \times \mathbf{K}_3} \cdot h^{\alpha - 2\alpha/k_0}$$

□

9 Piecewise Polynomial Functions and Convolutions

9.1 Continuous B -Splines Revisited

9.1.1 Scaled B -Splines Curves Associated to a Sequence

Let E be an \mathbb{R} -module and let $(\Gamma(s))_{s \in \mathbb{Z}}$ be a sequence with values in E . Let $r \in \mathbb{N}$. The uniform B -spline function with degree r and control points $(\Gamma(s))_{s \in \mathbb{Z}}$ is defined by:

$$\left(\mathcal{S}_{1,r}^{(0)}(\Gamma) \right) (t) = \sum_{s \in \mathbb{Z}} \Gamma(s) N_{s,r}(t)$$

where the $N_{s,r}$ function is the Cox De-Boor function (see Definition 5.3).

Definition 9.1 Let $l \in \mathbb{N}_+^*$ and $r \in \mathbb{N}$. First, for $s \in \mathbb{Z}$ we define the l -scaled B -spline function $N_{s,r,l}$ by:

$$N_{s,r,l}(t) = N_{s,r}\left(\frac{t}{l}\right) \tag{5}$$

Definition 9.2 The uniform l -scaled B -spline function with degree r associated to Γ is defined by:

$$\left(\mathcal{S}_{l,r}^{(0)}(\Gamma)\right)(t) = \sum_{s \in \mathbb{Z}} \Gamma(ls) N_{s,r,l}(t)$$

Note that this last function can be interpreted a regular uniform B -spline functions control points $(\Gamma(ls))_{s \in \mathbb{Z}}$ with difference between successive nodal points equal to l .

Definition 9.3 Let $\Phi : \mathbb{Z} \rightarrow E$ be a sequence, or $\Phi : \mathbb{R} \rightarrow E$ be a function. We denote by $\tau(\Phi)$ the right shift of Φ which to s associates $(\tau(\Phi))(s) = \Phi(s+1)$.

The following property follows immediately from the shift property in Lemma 5.1.

Property 9.1

$$\left(\mathcal{S}_{l,r}^{(0)}(\tau^l(\Gamma))\right)(t) = \left(\tau^l(\mathcal{S}_{l,r}^{(0)}(\Gamma))\right)(t)$$

9.1.2 Derivative of the B -splines Curves with Skipped Control Points

As far as the Cox de-Boor functions are concerned, we have the classical expression for the derivative:

$$N'_{s,r}(t) = (N_{s,r-1} - N_{s+1,r-1})$$

From the definition of $N_{s,r,l}(t)$ we get

$$N'_{s,r,l}(t) = \frac{1}{l} N'_{s,r}\left(\frac{t}{l}\right) = \frac{1}{l} (N_{s,r-1,l}(t) - N_{s+1,r-1,l}(t)) \quad (6)$$

Proposition 9.1 For $r \in \mathbb{N}^*$ we have:

$$\frac{d}{dt} \left(\mathcal{S}_{l,r}^{(0)}(\Gamma)\right)(t) = \frac{1}{l} \left(\mathcal{S}_{l,r-1}^{(0)}(\Gamma) - \mathcal{S}_{l,r-1}^{(0)}(\tau^{-l}(\Gamma))\right)$$

Proof.

$$\begin{aligned} \frac{d}{dt} \left(\mathcal{S}_{l,r}^{(0)}(\Gamma)\right)(t) &= \sum_{s \in \mathbb{Z}} \Gamma(ls) N'_{s,r,l}(t) \\ &= \sum_{s \in \mathbb{Z}} \Gamma(ls) \frac{1}{l} (N_{s,r-1,l}(t) - N_{s+1,r-1,l}(t)) \\ &= \sum_{s \in \mathbb{Z}} \Gamma(ls) \frac{1}{l} (N_{s,r-1,l}(t)) - \sum_{s \in \mathbb{Z}} \Gamma(l(s-1)) \frac{1}{l} (N_{s,r-1,l}(t)) \\ &= \frac{1}{l} \sum_{s \in \mathbb{Z}} (\Gamma(ls) - \tau^{-l}\Gamma(ls)) (N_{s,r-1,l}(t)) \end{aligned}$$

□

Definition 9.4 Let $\Phi : \mathbb{Z} \rightarrow E$ be a sequence, or $\Phi : \mathbb{R} \rightarrow E$ be a function. We denote

$$(\Delta_l^-(\Phi))(s) = \frac{1}{l} (\Phi(s) - \Phi(s-l)).$$

Notation 9.1 For $\omega \leq r-1$, we denote by $\mathcal{S}_{l,r}^{(\omega)}(\Gamma)$ the function defined on \mathbb{R} by:

$$\left(\mathcal{S}_{l,r}^{(\omega)}(\Gamma)\right)(t) = \frac{d^\omega}{dt^\omega} \left(\mathcal{S}_{l,r}^{(0)}(\Gamma)\right)(t)$$

With these notations, Proposition 9.1 can be reformulated as

$$\mathcal{S}_{l,r}^{(1)}(\Gamma) = \left(\Delta_l^- \left(\mathcal{S}_{l,r-1}^{(0)}(\Gamma) \right) \right)$$

Hence by induction, we get:

Proposition 9.2 For $\omega \leq r - 1$, we have

$$\mathcal{S}_{l,r}^{(\omega)}(\Gamma) = \left((\Delta_l^-)^\omega \left(\mathcal{S}_{l,r-\omega}^{(0)}(\Gamma) \right) \right)$$

9.1.3 Relationship between scaled B -spline and convolutions

Proposition 9.3 For $n \in \mathbb{Z}$, we have

$$\left(\mathcal{S}_{l,r}^{(0)}(\Gamma) \right) \left(l \left(n + \frac{r+1}{2} \right) \right) = \left(\mathbf{S}_{l,r}^{(0)} * \Gamma \right) (n)$$

Proof.

$$\begin{aligned} \left(\mathcal{S}_{l,r}^{(0)}(\Gamma) \right) \left(l \left(n + \frac{r+1}{2} \right) \right) &= \sum_{s \in \mathbb{Z}} \Gamma(ls) N_{s,r,l} \left(l \left(n + \frac{r+1}{2} \right) \right) \\ &= \sum_{s \in \mathbb{Z}} \Gamma(ls) N_{s,r} \left(n + \frac{r+1}{2} \right) \\ &= \sum_{s \in \mathbb{Z}} \Gamma(ls) N_{0,r} \left(\frac{r+1}{2} + n - s \right) \\ &= \sum_{s \in \mathbb{Z}} \Gamma(ls) \mathbf{S}_{l,r}^{(0)}(n - s) \\ &= \left(\mathbf{S}_{l,r}^{(0)} * \Gamma \right) (n) \end{aligned}$$

□

Notation 9.2 For $\omega \in \{0, \dots, r - 1\}$, we denote

$$\mathbf{S}_{l,r}^{(\omega)} = (\Delta_l^-)^\omega \left(\mathbf{S}_{l,r-\omega}^{(0)} \right)$$

Note that, from Proposition 3.1, the mask $\mathbf{S}_{l,r}^{(\omega)}$ is an ω -derivative mask.

Theorem 9.1 For $n \in \mathbb{Z}$ and $\omega \in \{0, \dots, r - 1\}$, we have

$$\left(\mathcal{S}_{l,r}^{(\omega)}(\Gamma) \right) \left(l \left(n + \frac{r+1}{2} \right) \right) = \left(\mathbf{S}_{l,r}^{(\omega)} * \Gamma \right) (n)$$

9.2 Bézier Curves with Skipped Control Points

9.2.1 Bézier Curves and Bernstein polynomials Basics

Definition 9.5 Let $r \in \mathbb{N}^*$. For $i \in 0, \dots, r$, we consider the \mathbb{R} -valued polynomial with degree r defined by

$$B_{i,r}(t) = \binom{r}{i} t^i (1-t)^{r-i}$$

These well-known polynomials are called the Bernstein polynomials with degree r .

In the sequel, unless otherwise specified, we shall say Bernstein polynomials or Bernstein functions as a shorthand for Bernstein polynomials. The Bernstein polynomials with degree r constitute a basis of the vector space of polynomials with degree less than or equal to r .

From the Pascal formula for binomial coefficients, we derive a similar formula about Bernstein polynomials:

$$B_{i,r}(t) = (1-t)B_{i,r-1}(t) + tB_{i-1,r-1}(t) \quad (7)$$

Definition 9.6 *Let E be an \mathbb{R} -module and let $\mathbf{P} = (P_0, \dots, P_{r-1})$ be a sequence of point in E . Let $B_{\mathbf{P}} : 1 \rightarrow E$ be the Bézier Curve with control points \mathbf{P} . Then we have:*

$$B_{\mathbf{P}}(t) = \sum_{i=0}^{r-1} P_i B_{i,r-1}(t)$$

We can compute the derivative of a Bézier Curve as follows. Let $B_{\mathbf{P}} : [0, 1] \rightarrow E$ be the Bézier Curve with control points P_0, \dots, P_{r-1} . Then we have:

$$B'_{\mathbf{P}}(t) = (r-1) \sum_{i=0}^{r-2} (P_{i+1} - P_i) B_{i,r-2}(t)$$

which is the Bézier curve with control points P'_1, \dots, P'_{n-1} , where $P'_i = (r-1)(P_{i+1} - P_i)$.

9.2.2 Scaled Bézier Function Associated to a Sequence

Definition 9.7 *Let $l \in \mathbb{N}_+^*$. We introduce the s -shifted l -scaled Bernstein polynomials by:*

$$B_{i,s,r,l}(t) = \begin{cases} B_{i,r}\left(\frac{t}{l} - s\right) & \text{if } \frac{t}{l} \in [s, s+1[\\ 0 & \text{otherwise.} \end{cases}$$

In other words, the value $B_{i,s,r,l}(t)$ can be non-zero only for $s = \lfloor \frac{t}{l} \rfloor$. In the remainder of this section, E denotes an \mathbb{R} -module and $(\Gamma(s))_{s \in \mathbb{Z}}$ is a sequence with values in E .

Definition 9.8 *Let $l \in \mathbb{N}_+^*$. The l -scaled (piecewise) Bézier curve with degree $r-1$ associated to Γ is defined over \mathbb{R} by:*

$$\left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) = \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B_{i,s,r,l}(t)$$

Note that in the previous definition, due to the definition of $B_{i,s,r,l}$, for a given value of i , only one value of s (namely $s = \lfloor \frac{t}{l} \rfloor + r - i$) contributes to the double sum $\left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t)$, so that, in fact, at most $r+1$ terms are non-zero.

Proposition 9.4 (De Casteljaou Property)

$$\left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) = \left(\left\lfloor \frac{t}{l} \right\rfloor + 1 - \frac{t}{l}\right) \left(\mathcal{B}_{l,r-1}^{(0)}(\tau^{-l}(\Gamma))\right)(t) + \left(\frac{t}{l} - \left\lfloor \frac{t}{l} \right\rfloor\right) \left(\mathcal{B}_{l,r-1}^{(0)}(\Gamma)\right)(t)$$

Proof.

$$\begin{aligned}
\left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) &= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B_{i,r}\left(\frac{t}{l} - s\right) \mathbb{1}_{[s, s+1[}\left(\frac{t}{l}\right) \\
&= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) \\
&\quad \cdot \left(\left(1 - \frac{t}{l} + s\right) B_{i,r-1}\left(\frac{t}{l} - s\right) + \left(\frac{t}{l} - s\right) B_{i-1,r-1}\left(\frac{t}{l} - s\right) \right) \\
&\quad \cdot \mathbb{1}_{[s, s+1[}\left(\frac{t}{l}\right)
\end{aligned}$$

The last equality follows from Equation 7. Taking into account that the only value of s for which $\frac{t}{l} \in [s, s+1[$, and then by changing the index i to $i+1$ in the sum, we get:

$$\begin{aligned}
\left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) &= \left(1 - \frac{t}{l} + \lfloor \frac{t}{l} \rfloor\right) \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B_{i,s,r-1,l}(t) + \\
&\quad \left(\frac{t}{l} - \lfloor \frac{t}{l} \rfloor\right) \sum_{s \in \mathbb{Z}} \sum_{i=-1}^{r-1} \Gamma(l(s+(i+1)-r)) B_{i,s,r-1,l}(t) \\
&= \left(1 - \frac{t}{l} + \lfloor \frac{t}{l} \rfloor\right) \sum_{s \in \mathbb{Z}} \sum_{i=0}^{r-1} \Gamma(l(s+(i-1)-(r-1))) B_{i,s,r-1,l}(t) + \\
&\quad \left(\frac{t}{l} - \lfloor \frac{t}{l} \rfloor\right) \sum_{s \in \mathbb{Z}} \sum_{i=0}^{r-1} \Gamma(l(s+i-(r-1))) B_{i,s,r-1,l}(t) \\
&= \left(1 - \frac{t}{l} + \lfloor \frac{t}{l} \rfloor\right) \left(\mathcal{B}_{l,r-1}^{(0)}(\tau^{-l}(\Gamma))\right)(t) + \left(\frac{t}{l} - \lfloor \frac{t}{l} \rfloor\right) \left(\mathcal{B}_{l,r-1}^{(0)}(\Gamma)\right)(t)
\end{aligned}$$

The indices $i = -1$ and $i = r$ yielding a zero term because out of range for the Bernstein polynomials. \square

9.2.3 Derivative of the Scaled Bézier Function

As far as Bernstein polynomials are concerned, we have the classical formula:

$$B'_{i,r} = r(B_{i-1,r-1} - B_{i,r-1})$$

We derive from this that, for $s = \lfloor t \rfloor$, we have

$$\frac{d}{dt} B_{i,s,r,l}(t) = \frac{1}{l} B'_{i,r}\left(\frac{t}{l} - s\right) = \frac{r}{l} (B_{i-1,s,r-1,l}(t) - B_{i,s,r-1,l}(t))$$

so that

Proposition 9.5 *For $r \geq 1$, the curve $\mathcal{B}_{l,r}^{(0)}(\Gamma)$ is C^{r-1} on \mathbb{R} and we have:*

$$\frac{d}{dt} \left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) = \frac{r}{l} \left(\mathcal{B}_{l,r-1}^{(0)}(\tau^l(\Gamma)) - \mathcal{B}_{l,r-1}^{(0)}(\Gamma)\right)(t)$$

Proof. First we prove the result for all $t \in \mathbb{R} \setminus \mathbb{Z}$, on which the curve $\mathcal{B}_{l,r}^{(0)}(\Gamma)$ is easily seen to be polynomial.

$$\begin{aligned}
\frac{d}{dt} \left(\mathcal{B}_{l,r}^{(0)}(\Gamma)\right)(t) &= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B'_{i,s,r,l}(t) \\
&= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) \frac{r}{l} (B_{i-1,s,r-1,l}(t) - B_{i,s,r-1,l}(t)) \\
&= \sum_{s \in \mathbb{Z}} \sum_{i=-1}^{r-1} \Gamma(l(s+(i+1)-r)) \frac{r}{l} (B_{i,s,r-1,l}(t)) - \\
&\quad \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) \frac{r}{l} (B_{i,s,r-1,l}(t)) \\
&= \sum_{s \in \mathbb{Z}} \sum_{i=0}^{r-1} \Gamma(l(s+1+i-r)) \frac{r}{l} (B_{i,s,r-1,l}(t)) - \\
&\quad \sum_{s \in \mathbb{Z}} \sum_{i=0}^{r-1} \Gamma(l(s+i-r)) \frac{r}{l} (B_{i,s,r-1,l}(t)) \\
&= \frac{r}{l} \left(\mathcal{B}_{l,r-1}^{(0)}(\tau^l(\Gamma)) - \mathcal{B}_{l,r-1}^{(0)}(\Gamma)(t)\right)
\end{aligned}$$

Now, as we can see that $\mathcal{B}_{l,1}^{(0)}(\Gamma)$ is C^0 on \mathbb{R} , so that the result follows by induction on $r \geq 2$. \square

Definition 9.9 Let $\Phi : \mathbb{Z} \rightarrow E$ be a sequence, or $\Phi : \mathbb{R} \rightarrow E$ be a function. We denote

$$(\Delta_l^+(\Phi))(s) = \frac{1}{l}(\Phi(s+l) - \Phi(s)).$$

Notation 9.3 For $\omega \in \{0, \dots, r-1\}$, we denote by $\mathcal{B}_{l,r}^{(\omega)}(\Gamma)$ the function on \mathbb{R} defined by

$$\mathcal{B}_{l,r}^{(\omega)}(\Gamma)(t) = \frac{d^\omega}{dt^\omega} \left(\mathcal{B}_{l,r-\omega}^{(0)}(\Gamma) \right) (t)$$

Therefore, Proposition 9.5 can be restated as:

$$\left(\mathcal{B}_{l,r}^{(1)}(\Gamma) \right) (t) = r \Delta_l^+ \left(\mathcal{B}_{l,r-1}^{(0)}(\Gamma) \right) (t)$$

The following immediately follows by induction:

Proposition 9.6 for $r \geq 1$ and $\omega \leq r-1$, we have

$$\left(\mathcal{B}_{l,r}^{(\omega)}(\Gamma) \right) (t) = \frac{r!}{(r-\omega)!} (\Delta_l^+)^{\omega} \left(\mathcal{B}_{l,r-\omega}^{(0)}(\Gamma) \right) (t)$$

9.2.4 Binomial Convolutions and Scaled Piecewise Bézier

Proposition 9.7

$$\left(\mathcal{B}_{l,r}^{(0)}(\Gamma) \right) \left(l \left(n + \frac{1}{2} \right) \right) = \left(\mathbf{B}_{l,r}^{(0)} * \Gamma \right) \left(l \left(n - \lfloor \frac{r}{2} \rfloor \right) \right)$$

Proof.

$$\begin{aligned} \left(\mathcal{B}_{l,r}^{(0)}(\Gamma) \right) \left(l \left(n + \frac{1}{2} \right) \right) &= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B_{i,s,r,l} \left(nl + \frac{l}{2} \right) \\ &= \sum_{s \in \mathbb{Z}} \sum_{i=0}^r \Gamma(l(s+i-r)) B_{i,r} \left(n + \frac{1}{2} - s \right) \mathbb{1}_{[s,s+1[} \left(n + \frac{1}{2} \right) \\ &= \sum_{i=0}^r \Gamma(l(n+i-r)) B_{i,r} \left(\frac{1}{2} \right) \\ &= \sum_{i=0}^r \Gamma(l(n-r+i)) \frac{1}{2^r} \binom{r}{i} \end{aligned}$$

We distinguish two cases depending on the parity of r :

First Case: If r is even. Say $r = 2m$ with $m \in \mathbb{N}$. We have:

$$\begin{aligned} \left(\mathbf{B}_{l,r}^{(0)} * \Gamma \right) (l(n-m)) &= \sum_{i \in \mathbb{Z}} \mathbf{B}_{l,2m}^{(0)}(li) \Gamma(l(n-m-i)) \\ &= \sum_{i \in \mathbb{Z}} \Gamma(l(n-2m+i)) \mathbf{B}_{l,2m}^{(0)}(l(m-i)) \\ &= \sum_{i=0}^{2m} \Gamma(l(n-r+i)) \frac{1}{2^{2m}} \binom{2m}{i} \end{aligned}$$

Second Case: If r is odd. Say $r = 2m-1$ with $m \in \mathbb{N}$.

$$\begin{aligned} \left(\mathbf{B}_{l,r}^{(0)} * \Gamma \right) (l(n-m+1)) &= \sum_{i \in \mathbb{Z}} \mathbf{B}_{l,2m}^{(0)}(li) \Gamma(l(n-m+1-i)) \\ &= \sum_{i \in \mathbb{Z}} \Gamma(l(n-2m+1+i)) \mathbf{B}_{l,2m}^{(0)}(l(m-i)) \\ &= \sum_{i=0}^{2m-1} \Gamma(l(n-(2m-1)+i)) \frac{1}{2^{2m-1}} \binom{2m-1}{i} \end{aligned}$$

□

Notation 9.4 For $\omega \in \{0, \dots, r-1\}$, we denote by $\mathbf{B}_{l,r}^{(\omega)}(\Gamma)$ the mask on \mathbb{R} defined by

$$\mathbf{B}_{l,r}^{(\omega)}(n) = (\Delta_l^+)^{\omega} (\mathbf{B}_{l,r-\omega})(n)$$

Note that this definition is coherent with Definition 7.1 and that $\mathbf{B}_{l,r}^{(\omega)}$ is an ω -derivative mask.

Theorem 9.2 For $r \geq 1$ and $\omega \leq r-1$, we have

$$\left(\mathcal{B}_{l,r}^{(\omega)}(\Gamma)\right) \left(l\left(n + \frac{1}{2}\right)\right) = \frac{r!}{(r-\omega)!} \left(\mathbf{B}_{l,r}^{(\omega)} * \Gamma\right) \left(l\left(n - \lfloor \frac{r}{2} \rfloor\right)\right)$$

Proof. The result follows by induction from Proposition 9.6 and Proposition 9.7. \square

10 Experiments

10.1 Smoothing

Figure 1 shows an example of smoothing a sampled function, with different models of noise of the sample values. Figure 2 shows the maximal error for smoothing with uniform noise, for different values of the resolution $\frac{1}{h}$.

We see that the Taylor-Optimal masks, which are optimal in the absence of noise, perform poorly for removing noise. The result of "smoothing" in this case is even more noisy than the original.

The best practical kernels seem to be that Binomial kernels. Although Proposition 5.1 indicates that the convergence order of the binomial masks with even or odd diameter is different, the practical performance for smoothing appears similar.

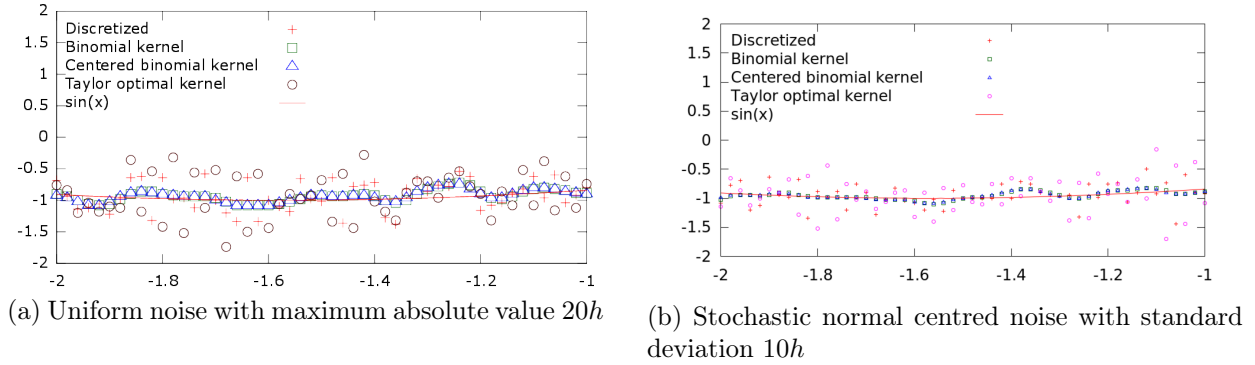
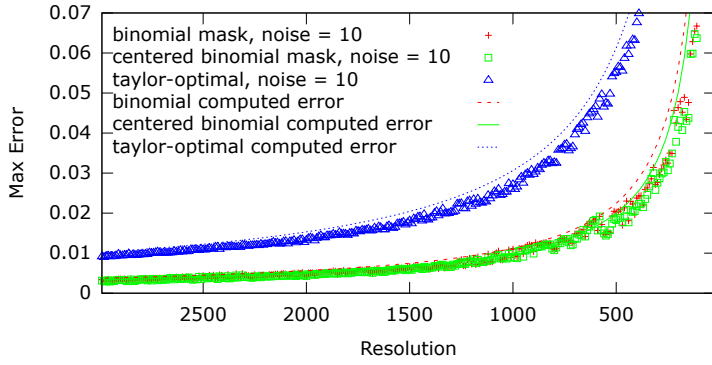
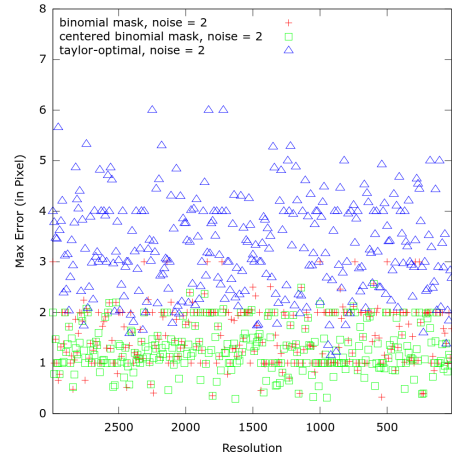


Figure 1: Smoothing operators applied to the sin function, digitization step $h = \frac{1}{50}$



(a) Maximum error in absolute value



(b) Maximum error with h as the unity

Figure 2: Smoothing operators applied to the sine function, uniform noise of maximum absolute value $2h$.

10.2 Derivatives

Figure 3 shows that binomial 1-derivative masks have a better practical error than the Taylor-optimal one in various cases with noise, which confirms the theoretical bounds.

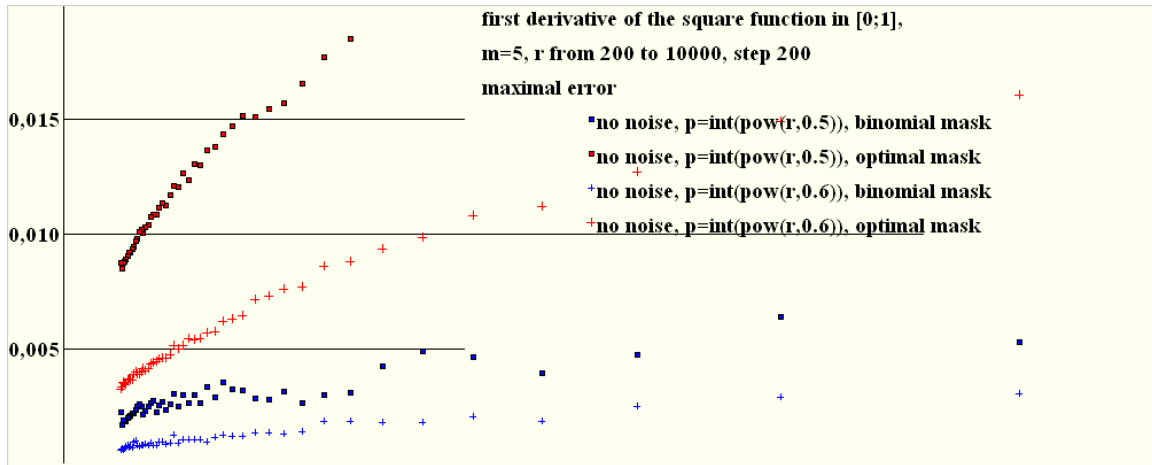


Figure 3: Maximal error 1-derivative operators applied to the square function, for different values of the resolution $\frac{1}{h}$ from 200 to 1000. The results are for a **uniform noise** with maximum absolute value $\frac{1}{4}h^{0.8}$ (i.e. $\alpha = 0.8$). Mask cardinality $r = 11$. Skipping step for $l = \lfloor h^{-0.5} \rfloor$ and for $l = \lfloor h^{-0.6} \rfloor$

Figure 4, also confirms the result that the Taylor-Optimal 1-derivative masks perform better in the absence of noise, but similar, or even worse in the presence of very moderate noise (close to round-off quantisation errors). We see that the performance of the Binomial masks have a closer dependency to the resolution, but still have a broad spectrum of good performance.

We also see on both Figure 3 and Figure 4 that a good estimation can be obtained by a mask using only 11 (possibly noisy) values. This shows the strength, as far as computational

complexity is concerned, of the method with skipping steps compared to using contiguous samples.

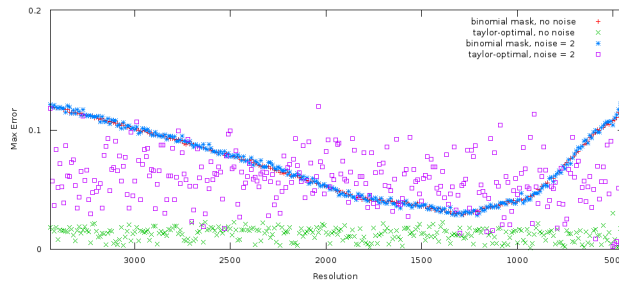


Figure 4: Maximal error for 1–derivative operators applied to the sin function, uniform noise of maximum absolute value $2h$, mask cardinality $r = 11$ and $l = 100$

Figure 5 shows that, also with only 40 points, the second derivative can be estimated with good precision in the round-off quantization value errors case. The figure also shows the influence of the skipping step. Below some threshold power of h , the method just does not converge.

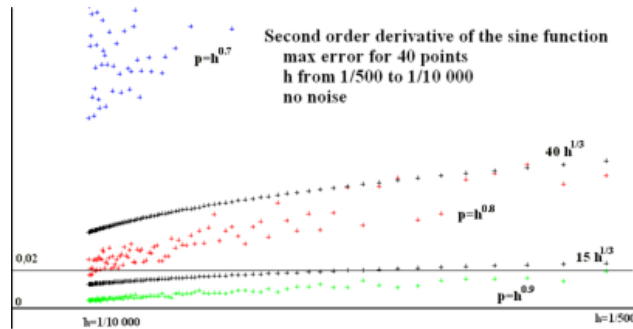


Figure 5: Maximal error of 2–derivative operators with binomial skipping masks, on the sin function, with only round-off quantization value errors. All convolution values are computed with 40 samples, for different values $l = h^{0.7}$, $l = 40h^{1/3}$, $l = h^{0.8}$, $l = 15h^{1/3}$, $l = h^{0.9}$

10.3 Curvature

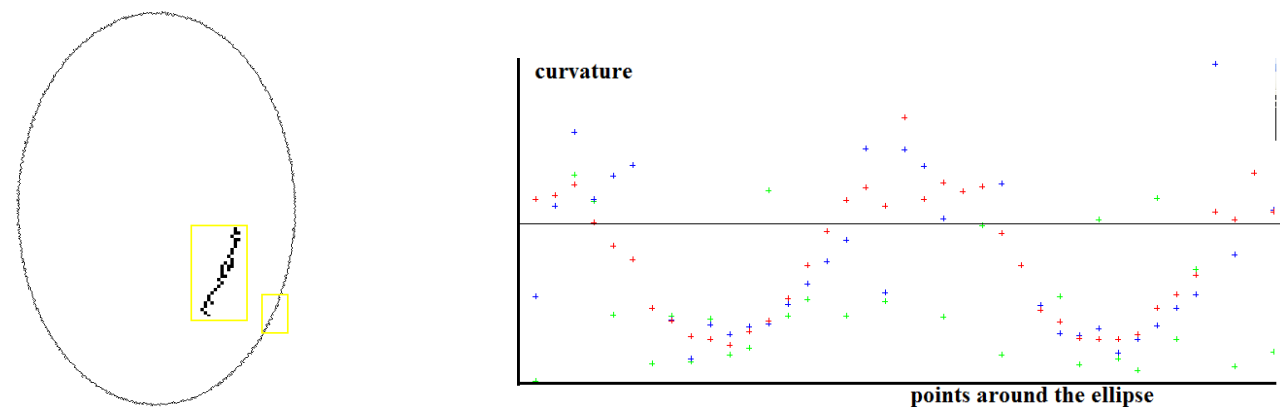


Figure 6: Digital curvature around a noisy digitization of an ellipse, uniform noise of maximum absolute value $3h$, $h = \frac{1}{500}$, Taylor-optimal mask, various values for l

Acknowledgement

We thank Geoffrey Giry for his contribution in generating the figures 1, 2, and 4.

11 Conclusion

This work is a milestone in our work in progress to define a new paradigm for numerical analysis and calculus, to avoid the shortcomings of floating point computations. It can be easily extended to any dimension by tensor product, which leads to convergence results by application of the one-dimensional case.

The interpretation of digital convolutions using smooth curves can be generalized to many convolution kernels, including the Gaussian kernels, for which the smooth curve can be chosen C^∞ . Even though we are primarily interested in integer-only or exact computations, this is an interesting limit case, which can be of interest to many scientists using more traditional approaches to numerical analysis and functional analysis.

We will build on the general notion of a digital derivative to define other tools for digital analysis, such as integrals, partial differential equations, convex optimization, harmonic analysis, and non linear optimization. Some of this work is already in progress as we write these lines.

We also plan to develop a dedicated library, for efficient integer-only computations. The library should be generic and support many formats for numbers, including long-integers and floating points. More experiments on n -dimensional derivative and curvature estimation will be performed in this process. Even-though long-integers come with an overhead, we believe that exact calculations will prevail over cheap approximations in the long run.

A Choosing a Mask

A.1 Criteria for Choosing a Mask

Having read all of the paper until the conclusion, the reader may still be wondering: "Why bother with so many masks?". In fact, there is no single bullet rule for choosing a mask. For

now, any user of our methods has to assess what we call “business dependant” or “application dependant” hypothesis, about available resolutions and noise model.

If you have low resolution and low noise (a few pixels), you should probably use the Taylor Optimal masks (see Figure 4). If you have a fair resolution and significant noise, you will probably be better off with a Binomial mask.

The B -spline masks have not yet been completely assessed, and they are here for their potential theoretical interest, which deserves the proofs of multigrid convergence to be included.

In the next section, we attempt to provide finer criteria to choose masks for your favourite application.

A.2 Moments of The Main Considered Masks

In Lemma 4.1, an upper bound is given for the sampling error $|ES_\omega(f, h, \Gamma, \mathbf{u}, n)|$, as a function of the infinite norm of $f^{(k_0)}$, the digitization step, and a constant which depends on \mathbf{u} : $\sum_{i \in \mathbb{Z}} |i^{k_0} u(i)|$. For $k_0 = 1 + \rho$, where ρ is the convergence order of the digital ω -derivative mask \mathbf{u} , we call this last constant, the *moment* of the mask.

This constant gives an indication on the speed of convergence (main asymptotic term) of the sampling error the digital ω -derivative estimation given by \mathbf{u} for $C^{1+\rho}$ functions. We can also consider this constant for $k_0 \leq \rho$ to assess the quality of estimation on functions which are not $C^{1+\rho}$.

Similarly, in Lemma 4.2 and in the proof of Theorem 6.1, the main asymptotic term in the error on values $|EV_\omega(f, h, \Gamma, \mathbf{u}, n)|$ depends on a constant $\sum_{i \in \mathbb{Z}} |u(i)|$. We also asses this values for some masks.

As we advocate the use of masks of small support, we give here schedules for the coefficients values.

A.3 Smoothing Masks

m	3	4	5	6	7	8	9
$\frac{1}{(2m)!} \sum_{i \in \mathbb{Z}} i^{2m} \mathbf{O}_{2m}^{(0)}(i) $	0.157	0.089	0.051	0.029	0.017	0.010	0.0053
$\frac{1}{(2m)!} \sum_{i \in \mathbb{Z}} \mathbf{O}_{2m}^{(0)}(i) $	3.2	3.66	4.07	4.44	4.78	5.10	5.40

Table 1: Moments of the mask $\mathbf{O}_{2m}^{(0)}$

m	3	4	5	6	7	8	9
$\sum_{i \in \mathbb{Z}} i \mathbf{B}_{2m-1}^{(0)}(i) $	0.94	1.10	1.24	1.36	1.47	1.58	1.67

Table 2: Moments of the mask $\mathbf{B}_{2m-1}^{(0)}$

For $\mathbf{B}_{2m}^{(0)}$ we have the simple formula $\frac{1}{2} \sum_{i \in \mathbb{Z}} i^2 |\mathbf{B}_{2m}^{(0)}(i)| = \frac{m}{4}$ (see the proof of Proposition 5.1).

$r2$		3	4	5	6	7	8
$\frac{1}{2} \sum_{i \in \mathbb{Z}} i^2 \mathbf{S}_r^{(0)}(i) $	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{125}{600}$	$\frac{1}{4}$	$\frac{175}{600}$	$\frac{1}{6}$	$\frac{3}{8}$

Table 3: Moments of the mask $\mathbf{S}_r^{(0)}$

A.4 First Order Derivative Masks

In practice, m will be 3 to 7. Hence we first give a table of values for such m .

m	3	4	5	6	7	8	9
k							
0	$\frac{1.84}{l}$	$\frac{2.09}{l}$	$\frac{2.29}{l}$	$\frac{2.45}{l}$	$\frac{2.60}{l}$	$\frac{2.72}{l}$	$\frac{2.83}{l}$
1	2.2	2.66	3.07	3.44	3.78	4.10	4.40
2	$3.0l$	$4.0l$	$5.0l$	$6.0l$	$7.0l$	$8.0l$	$9.0l$
3	$4.8l^2$	$7.32l^2$	$10.2l^2$	$13.3l^2$	$16.8l^2$	$20.4l^2$	$24.3l^2$
4	$9l^3$	$16l^3$	$25l^3$	$36l^3$	$49l^3$	$64l^3$	$81l^2$

Table 4: $\sum_{i \in \mathbb{Z}} |i^k \mathbf{O}_{l,m}^{(1)}(i)|$

m	3	4	5	6	7	8	9
k							
0	$\frac{0.625}{l}$	$\frac{0.547}{l}$	$\frac{0.493}{l}$	$\frac{0.452}{l}$	$\frac{0.419}{l}$	$\frac{0.393}{l}$	$\frac{0.371}{l}$
1	1	1	1	1	1	1	1
2	$1.875l$	$2.188l$	$2.461l$	$2.707p$	$2.932p$	$3.142p$	$3.338p$
3	$4.0l^2$	$5.5l^2$	$7l^2$	$8.5l^2$	$10.0l^2$	$11.5l^2$	$13l^2$

Table 5: $\sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{l,2m-1}^{(1)}(i)|$

r	2	3	4	5	6	7	8
k							
0	$\frac{0.875}{l}$	$\frac{0.959}{l}$	$\frac{0.797}{l}$	$\frac{0.877}{l}$	$\frac{0.740}{l}$	$\frac{0.806}{l}$	$\frac{0.694}{l}$
1	1	1	1	1	1	1	1
2	$2.125l$	$1.563l$	$2.209l$	$1.687p$	$2.92p$	$1.806p$	$2.375p$
3	$4.75l^2$	$2.75l^2$	$5.25l^2$	$3.25l^2$	$5.75l^2$	$3.75l^2$	$6.25l^2$

Table 6: $\sum_{i \in \mathbb{Z}} |i^k \mathbf{S}_{l,r}^{(1)}(i)|$

In order to compare both masks, let us consider a C^2 function f and $m = 5$.

In the binomial case, the error bound is $\frac{7}{6} \|f^{(3)}\|_{\infty} (l(h)h)^2 + \frac{K}{2} \frac{h^{\alpha-1}}{l(h)}$.

In the optimal case, the error bound is $\frac{11}{6} \|f^{(3)}\|_{\infty} (l(h)h)^2 + \frac{5K}{2} \frac{h^{\alpha-1}}{l(h)}$ which is worse.

For a C^3 function, the error bound remains the same in the binomial case, but in the optimal case, it become $\frac{25}{24} \|f^{(4)}\|_{\infty} (l(h)h)^3 + \frac{5K}{2} \frac{h^{\alpha-1}}{l(h)}$.

The binomial mask seems to be competitive in the generic case. Hence, we give here general formulas for this case:

Proposition A.1 $\sum_{i \in \mathbb{Z}} |\mathbf{B}_{l,2m-1}^{(1)}(i)| = \frac{1}{l^{2m-2}} \binom{2m-1}{m-1}$

$$\sum_{i \in \mathbb{Z}} |i \mathbf{B}_{l,2m-1}^{(1)}(i)| = 1$$

$$\sum_{i \in \mathbb{Z}} |i^2 \mathbf{B}_{l,2m-1}^{(1)}(i)| = \frac{(2m-1)l}{2^{2m-2}} \binom{2m-2}{m-1}$$

$$\sum_{i \in \mathbb{Z}} |i^3 \mathbf{B}_{l,2m-1}^{(1)}(i)| = \frac{3m-1}{2} l^2$$

Proof. Let us notice that for odd k , $|i^k \mathbf{B}_{l,2m-1}(i)| = i^k \mathbf{B}_{l,2m-1}(i)$, hence $\sum_{i \in \mathbb{Z}} |i \mathbf{B}_{l,2m-1}(i)| = 1$ and $\sum_{i \in \mathbb{Z}} |i^3 \mathbf{B}_{l,2m-1}(i)| = \frac{3m-1}{2} p^2$.

Now for even k , we have

$$\sum_{i < 0} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-1}} \sum_{j=-m}^{j=-1} |j^k| \left(\binom{2m-1}{m-1-j} - \binom{2m-1}{m-j} \right)$$

$$\sum_{i < 0} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-1}} \sum_{j=1}^{j=m} j^k \left(\binom{2m-1}{m-1+j} - \binom{2m-1}{m+j} \right)$$

$$\sum_{i < 0} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-1}} \sum_{j=1}^{j=m} j^k \left(\binom{2m-1}{m-j} - \binom{2m-1}{m-1-j} \right)$$

$$\text{hence } \sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-2}} \sum_{j=1}^{j=m} j^k \left(\binom{2m-1}{m-j} - \binom{2m-1}{m-1-j} \right).$$

For $k = 0$, we have $\sum_{i \in \mathbb{Z}} |\mathbf{B}_{l,m}^{(1)}(i)| = \frac{1}{l^{2m-2}} \binom{2m-1}{m-1}$.

For $k \geq 2$, we have

$$\sum_{j=1}^{j=m} j^k \left(\binom{2m-1}{m-j} - \binom{2m-1}{m-1-j} \right) = \sum_{j=1}^{j=m} j^k \binom{2m-1}{m-j} - \sum_{j=1}^{j=m} j^k \binom{2m-1}{m-1-j}$$

and

$$\sum_{j=1}^{j=m} j^k \binom{2m-1}{m-1-j} = \sum_{j=2}^{j=m+1} (j-1)^k \binom{2m-1}{m-j},$$

$$\text{hence } \sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-2}} \sum_{j=1}^{j=m} (j^k - (j-1)^k) \binom{2m-1}{m-j}$$

For $k = 2$, we have

$$\sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{l,m}^{(1)}(i)| = \frac{l^{k-1}}{2^{2m-2}} \sum_{j=1}^{j=m} (2j-1) \binom{2m-1}{m-j} = \frac{p^{k-1}}{2^{2m-2}} (2m-1) \binom{2m-2}{m-1} \square$$

A.5 Second Order Derivatives

m	3	4	5	6	7	8	9
$\sum_{i \in \mathbb{Z}} \mathbf{B}_{2m-1}^{(2)}(i) $	0.63	0.44	0.38	0.33	0.28	0.0.25	0.22
$\sum_{i \in \mathbb{Z}} i^2 \mathbf{B}_{2m-1}^{(2)}(i) $	2.32	2.22	2.29	2.30	2.28	2.27	2.25
$\sum_{i \in \mathbb{Z}} i^3 \mathbf{B}_{2m-1}^{(2)}(i) $	6.57	7.33	8.25	9.00	9.64	10.23	10.77

Table 7: $\sum_{i \in \mathbb{Z}} |i^k \mathbf{B}_{2m-1}^{(2)}(i)|$

References

- [AM00] Götz Alefeld and Günter Mayer. Interval analysis: theory and applications. *Journal of Computational and Applied Mathematics*, 121(12):421 – 464, 2000.
- [CLL13] David Coeurjolly, Jacques-Olivier Lachaud, and Jérémy Levallois. Integral based curvature estimators in digital geometry. In *Discrete Geometry for Computer Imagery, DGCI 2013, Seville, Spain, March 20-22, 2013. Proceedings*, volume 7749 of *Lecture Notes in Computer Science*, pages 215–227. Springer, 2013.
- [CLL14] David Coeurjolly, Jacques-Olivier Lachaud, and Jérémy Levallois. Multigrid convergent principal curvature estimators in digital geometry. *Computer Vision and Image Understanding*, 129:27–41, 2014.
- [CLMT15] Louis Cuel, Jacques-Olivier Lachaud, Quentin Mérigot, and Boris Thibert. Robust geometry estimation using the generalized voronoi covariance measure. *SIAM J. Imaging Sciences*, 8(2):1293–1314, 2015.
- [Coh03] Joel S. Cohen. *Computer Algebra and Symbolic Computation: Mathematical Methods*. AK Peters, Ltd, Natick Massachusetts, 2003.
- [dVL09] François de Vieilleville and Jacques-Olivier Lachaud. Comparison and improvement of tangent estimators on digital curves. *Pattern Recognition*, 42(8):1693–1707, 2009.
- [EM14] Henri-Alex Esbelin and Rémy Malgouyres. Taylor optimal kernel for derivative estimation. In *Discrete Geometry for Computer Imagery - 18th IAPR International Conference, DGCI 2014, Siena, Italy, September 10-12, 2014. Proceedings*, volume 8668 of *Lecture Notes in Computer Science*, pages 384–395. Springer, 2014.
- [EMC11] Henri-Alex Esbelin, Rémy Malgouyres, and Colin Cartade. Convergence of binomial-based derivative estimation for c^2 noisy discretized curves. *Theor. Comput. Sci.*, 412(36):4805–4813, 2011.
- [Flo05] Michael S. Floater. On the convergence of derivatives of bernstein approximation. *Journal of Approximation Theory*, 134(1):130–135, 2005.

- [Flo08] Michael S. Floater. On the deviation of a parametric cubic spline interpolant from its data polygon. *Computer Aided Geometric Design*, 25(3):148–156, 2008.
- [FM09] Sébastien Fourey and Rémy Malgouyres. Normals estimation for digital surfaces based on convolutions. *Computers & Graphics*, 33:2–10, 2009.
- [GMES13] Damien Gonzalez, Rémy Malgouyres, Henri-Alex Esbelin, and Chafik Samir. Convergence of level-wise convolution differential estimators. In *DGCI*, volume 7749 of *Lecture Notes in Computer Science*, pages 335–346. Springer, 2013.
- [KL09] Bertrand Kerautret and Jacques-Olivier Lachaud. Curvature estimation along noisy digital contours by approximate global optimization. *Pattern Recognition*, 42(10):2265–2278, 2009.
- [KLN08] Bertrand Kerautret, Jacques-Olivier Lachaud, and Benoît Naegel. Comparison of discrete curvature estimators and application to corner detection. In *Advances in Visual Computing, 4th International Symposium, ISVC 2008, Las Vegas, NV, USA, December 1-3, 2008. Proceedings, Part I*, volume 5358 of *Lecture Notes in Computer Science*, pages 710–719. Springer, 2008.
- [KSW05] Ron Kimmel, Nir A. Sochen, and Joachim Weickert, editors. *Scale Space and PDE Methods in Computer Vision, 5th International Conference, Scale-Space 2005, Hofgeismar, Germany, April 7-9, 2005, Proceedings*, volume 3459 of *Lecture Notes in Computer Science*. Springer, 2005.
- [LCL14] Jérémy Levallois, David Coeurjolly, and Jacques-Olivier Lachaud. Parameter-free and multigrid convergent digital curvature estimators. In *Discrete Geometry for Computer Imagery - 18th IAPR International Conference, DGCI 2014, Siena, Italy, September 10-12, 2014. Proceedings*, pages 162–175, 2014.
- [Lin90] Tony Lindeberg. Scale-space for discrete signals. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(3):234–254, 1990.
- [LVdV05] Jacques-Olivier Lachaud, Anne Vialard, and François de Vieilleville. Analysis and comparative evaluation of discrete tangent estimators. In *Discrete Geometry for Computer Imagery, 12th International Conference, DGCI 2005, Poitiers, France, April 13-15, 2005, Proceedings*, volume 3429 of *Lecture Notes in Computer Science*, pages 240–251. Springer, 2005.
- [LVdV07] Jacques-Olivier Lachaud, Anne Vialard, and François de Vieilleville. Fast, accurate and convergent tangent estimation on digital contours. *Image Vision Comput.*, 25(10):1572–1587, 2007.
- [MB14] Loïc Mazo and Etienne Baudrier. About multigrid convergence of some length estimators. In *Discrete Geometry for Computer Imagery - 18th IAPR International Conference, DGCI 2014, Siena, Italy, September 10-12, 2014. Proceedings*, volume 8668 of *Lecture Notes in Computer Science*, pages 214–225. Springer, 2014.
- [MBF08] R. Malgouyres, F. Brunet, and S. Fourey. Binomial convolutions and derivatives estimation from noisy discretizations. In *Discrete Geometry for Computer Imagery, LNCS*, volume 4992, pages 370–379, 2008.

- [Moo79] R. E. Moore. *Methods and Applications of Interval Analysis*. Studies in Applied and Numerical Mathematics. SIAM, 1979.
- [PG11] Laurent Provot and Yan Gérard. Estimation of the derivatives of a digital function with a convergent bounded error. In *Discrete Geometry for Computer Imagery - 16th IAPR International Conference, DGCI 2011, Nancy, France, April 6-8, 2011. Proceedings*, volume 6607 of *Lecture Notes in Computer Science*, pages 284–295. Springer, 2011.
- [Vui87] J. Vuillemin. Exact real computer arithmetic with continued fractions. Technical Report RR-0760, INRIA, Nov 1987.