Computer Aided Geometric Design 29 (2012) 150–161

Contents lists available at SciVerse ScienceDirect

Computer Aided Geometric Design

www.elsevier.com/locate/cagd

L-system specification of knot-insertion rules for non-uniform B-spline subdivision ☆

V. Nivoliers [a,*], C. Gérot [b,**], V. Ostromoukhov [c], N.F. Stewart [d]

[a] *ALICE, Inria, France*
[b] *GIPSA-lab, Grenoble, France*
[c] *LIRIS, CNRS, France*
[d] *LIGUM, Université de Montréal, Canada*

ARTICLE INFO

Article history:
Received 17 March 2011
Received in revised form 24 November 2011
Accepted 29 November 2011
Available online 7 December 2011

Keywords:
L-system
Subdivision
Non-uniform

ABSTRACT

Subdivision schemes are based on a hierarchy of knot grids in parameter space. A univariate grid hierarchy is regular if all knots are equidistant on each level, and irregular otherwise. We use L-systems to design a wide class of systematically described irregular grid hierarchies. Furthermore, we give sufficient conditions on the L-system which guarantee that the subdivision scheme, based on the non-uniform B-spline of degree d defined on the initial knot grid, is uniformly convergent. If n is the number of symbols in the alphabet of the L-system, this subdivision scheme is defined with a finite set of masks (at most n^{d+1}) which does not depend on the subdivision step. We provide an implementation of such schemes which is available as a worksheet for Sage software.

1. Introduction

The representation of smooth objects (functions, curves or surfaces) is a fundamental problem in computer graphics and geometric modelling. One efficient representation consists of parametrising the object with a spline expressed in terms of a B-spline basis. Each polynomial piece of the spline is defined on an interval (or rectangular domain) whose vertices are called knots; the B-spline basis allows us to associate with the knots in parameter space a set of control points in object space, whose positions influence only locally the shape of the object, thus allowing intuitive control.

Subdivision schemes provide an efficient way to draw such smooth objects from given nets of control points and knots, and a given degree (or bi-degree) for the B-spline basis. These schemes consist of inserting new knots and successively computing new control nets, each defining the same smooth object. If the successively inserted knots are dense enough then the control net converges to the smooth object. In practice, a few steps are enough to reach the resolution of a computer screen.

When the knots define a regular grid, often assumed without loss of generality to be $\mathbb{Z}^N$, where N is the dimension of the object, the B-spline basis is uniform. If the new knots are inserted at midpoints, the new B-spline basis remains uniform and the uniform subdivision rules which define the new control net are the same convex combinations, whatever the interval length: they depend only on N and on the degree of the B-spline basis. Such uniform subdivision schemes are the most commonly used.

☆ This paper has been recommended for acceptance by Hartmut Prautzsch.

* Corresponding author.
** Principal corresponding author.
E-mail address: Cedric.Gerot@gipsa-lab.grenoble-inp.fr (C. Gérot).

0167-8396/$ – see front matter
doi:10.1016/j.cagd.2011.11.004

In some cases, however, non-uniform subdivision schemes are needed, either to represent objects which cannot be parametrised with a uniform spline or to deal with knots that do not lie on $2^{-k}\mathbb{Z}^N$ (for example in the context of multi-resolution analysis with irregular samples). We are interested in the non-uniform case: we propose subdivision schemes for B-spline parametrised curves with irregular but controlled knot intervals. In this paper we restrict our attention to univariate schemes ($N = 1$).

Algorithms for inserting one (Boehm, 1980) or several knots (Cohen et al., 1980; Barry and Zhu, 1992) into intervals separating non-equally spaced knots have been known for decades. They lead to subdivision schemes with knot sequences that are possibly completely irregular, and with as many subdivision processes as there are knot intervals in all of the successive subdivision steps.

The method of Non-uniform Recursive Subdivision Surfaces proposed by Sederberg et al. (1998) starts with irregular knot intervals and inserts, at each step, one knot at the midpoint of each interval. This is, therefore, bisection with rules whose coefficients are written as functions of successive interval length, which consequently have to be computed at each subdivision step, for each vertex.

Recently, algorithms which adapt the efficient refine-and-smooth factorisation, proposed by Lane and Riesenfeld (1980) for uniform B-splines, to the non-uniform B-splines, have been proposed (Schaefer and Goldman, 2009; Cashman et al., 2009b). These algorithms define the subdivision as a single process but they require that exactly one knot be inserted in every interval.

The algorithm proposed by Schaefer and Goldman (2009) can be generalised to the insertion of more than one knot in every interval as long as the number of inserted knots is the same whatever the interval within a subdivision step. Cashman et al. (2009a) permit the omission of subdivision of some intervals, but, in order to preserve the locality of the smoothing steps, they do not support the insertion of more than one knot per interval. Whichever algorithm is considered, the positions of inserted knots have to be explicitly given at each subdivision step, necessitating the computation of new coefficients for the rules used at each subdivision step.

Another approach, suggested by Goldman and Warren (1993), considers an affine relationship between the knots of the initial grid and inserts one knot at a constant barycentric position between each pair of adjacent old knots to preserve the affine relationship. The same insertion rule is used at each step, and everywhere in the domain.

We propose a framework to describe a wide class of non-uniform subdivision schemes with knot sequences more irregular than uniform or affine bisection (or trisection, etc.), but with a small set of subdivision processes which does not depend on the subdivision step, and using fixed coefficients that can be computed in advance.

The proposed framework uses context-free L-systems (Herman et al., 1974) to describe the sequences of irregular knot intervals. The specification of uniform schemes for curves (Prusinkiewicz et al., 2002) and surfaces (Velho, 2003), using L-systems, with the goal of easing implementation, has been done previously, by translating known subdivision procedures into the rules of a context-sensitive grammar.

In Section 2 we show that an L-system is ideal for the design of interval subdivision descriptors which are based on lengths of knot intervals rather than on the position of the knots themselves, and which can be used to define non-uniform B-spline subdivision schemes. We also give details on the computation of the subdivision masks. Not every L-system, however, can be used as such an interval subdivision descriptor. For this reason we introduce in Section 3 the concept of a valid L-system, which yields in particular a convergent non-uniform B-spline subdivision scheme, and we introduce sufficient conditions on the rules for an L-system to be valid.

2. Non-uniform subdivision scheme from L-systems

When knots are inserted, the new control points can be expressed as convex combinations of old control points. These combinations depend only on the degree of the B-spline and the two successive sequences of knots. The difference between B-spline subdivision schemes and knot-insertion algorithms is that subdivision inserts many knots at the same time. The cost of these insertions is reduced due to the fact that the necessary data is shared in the computation of the new control points. In uniform subdivision, this sharing is observable in the subdivision mask which collects, for a given control point, the coefficients of its contributions to the subdivided control points. For a subclass of non-uniform subdivision where at most one knot is inserted in every interval, Cashman et al. (2009a) propose a refine-and-smooth implementation which is another way of sharing data within the subdivision process.

Our aim is to give a different subclass of non-uniform B-spline subdivision which allows us to define a finite set of masks that remains the same whatever the subdivision step. Our framework remains non-uniform in the sense that the subdivision rules are not the same everywhere within one subdivision step and may insert different numbers of knots in each interval.

2.1. The L-system as an interval subdivision descriptor

Using an L-system, we describe the subdivision in terms of splitting knot intervals, and we base our formalism on the lengths of the intervals between the knots rather than the position of the knots themselves.

2.1.1. L-systems

A Lindenmayer system or L-system is a particular kind of *rewriting system* where the rules are applied greedily to rewrite as many symbols as possible at each step, whereas usually only one symbol is rewritten at a time. These systems have

mainly been used to model plants, other organisms, or such things as the buildings in a city, and also to generate images of a wide class of fractals with a geometric interpretation of strings based on the notion of a LOGO-style turtle (Prusinkiewicz and Hanan, 1989; Herman et al., 1974). Formally, they are defined by a grammar, which is a tuple $G = (\Sigma, \Pi, \alpha)$ where

- Σ is the alphabet of symbols,
- Π is the set of rewriting rules,
- α is the initial configuration (axiom).

Let Σ^* be the set of all words over Σ, and let Σ^+ be the set of all non-empty words over Σ. The axiom α is assumed to be a member of Σ^+. A rule $\pi \in \Pi$ is a mapping of a symbol $s \in \Sigma$ to a word $s' \in \Sigma^*$, i.e., $\pi : \Sigma \mapsto \Sigma^*$. Many extensions of L-systems have been defined, including context-sensitive or parametrised L-systems.

The concept of subdivision has already been linked with context-sensitive parametrised L-systems in Prusinkiewicz et al. (2002), Velho (2003). Prusinkiewicz addressed the univariate uniform case, by using a single symbol P in the grammar to represent a control point, with a parameter v representing its coordinates. The control polygon is then represented by a word on this symbol, and the rules encode the subdivision stencils. For example, Chaikin's subdivision for closed curves is described by a single context-sensitive rule rewriting P as PP with some convenient parameter defining the coordinates, where the context of $P(v)$ is made up of its left neighbour $P(v_l)$ and its right neighbour $P(v_r)$:

$$P(v_l)\langle P(v)\rangle P(v_r) \to P\left(\frac{1}{4}v_l + \frac{3}{4}v\right)P\left(\frac{3}{4}v + \frac{1}{4}v_r\right).$$

Velho extended this model to the bivariate case, replacing the concept of words in Σ^* by a concept of mesh with half-edge structure, and rules on the mesh.

Our formalism describes the subdivision in terms of interval decomposition. Each symbol labels an interval between two knots, and the context-free L-system rules describe the subdivision of these intervals resulting from the insertion of new knots.

2.1.2. Principle

The alphabet of our L-system is a set of symbols $\{A_i\}_{i=1}^n$ which are labels associated with knot intervals. Each rule describes how one labelled interval is split into an ordered sequence of labelled knot intervals:

$$A_i \to A_{i_1} \dots A_{i_r} \tag{1}$$

where $r \geqslant 1$ and each i_j, $j \in \{1, \dots, r\}$ is an element of $\{1, \dots, n\}$. The axiom of our L-system is the word defined by the symbols attached to the consecutive knot intervals in the original knot setting. To complete the definition of such an L-system as an interval subdivision descriptor, a length a_i^s must be associated with each symbol A_i and each subdivision step s such that (1) yields

$$a_i^s = \sum_{j=1}^{r} a_{i_j}^{s+1}. \tag{2}$$

One subdivision step corresponds to rewriting the word of the interval sequence, by replacing all symbols simultaneously by their split versions. Applied to the definition of a subdivision scheme, this corresponds to knot-insertions in parameter space. In order to define a set of masks that does not depend on the subdivision step, we define one length a_i for each symbol A_i and a ratio ρ that depends neither on the interval nor on the subdivision step, and which is used to scale down all of the interval lengths. Thus, for all labels A_i and for all subdivision steps s, $a_i^s = a_i/\rho^s$ and (2) becomes

$$a_i = \sum_{j=1}^{r} \frac{1}{\rho} a_{i_j}. \tag{3}$$

Moreover, in order to get a dense set of knots when the interval subdivision is applied repeatedly, which will be used in Section 3.1 to show the convergence of the subdivision scheme, we should have $\rho > 1$.

With this formalism, a uniform bisecting subdivision scheme defines one label A and one rule $A \to AA$ with any positive length a and the ratio $\rho = 2$, whereas the non-uniform subdivision schemes proposed by Sederberg et al. (1998) define, in general, as many labels A_i, as many rules $A_i \to A_i A_i$, and as many lengths a_i as edges in the given control polygon, with the unique ratio $\rho = 2$.

But not every set of rules allows the existence of lengths $\{a_i\}$ and ratio ρ satisfying (3). Consequently, we introduce in Section 3 the concept of a valid system. Then, we give sufficient conditions for the system to be valid and show how appropriate lengths $\{a_i\}$ and ratio $\rho > 1$ can be computed automatically. Note that in our framework, we define knot-interval lengths from the rules and not the rules from given lengths of intervals.

In order to show the central idea, we first explain how to compute the finite set of masks from an L-system.

2.2. Subdivision masks from interval subdivision

Describing knot-interval decomposition with an L-system allows us to generalise the concept of mask to non-uniform subdivision schemes. The set of subdivision masks is computed from the interval lengths $\{a_i\}$, the ratio ρ and the degree d of the spline.

Let $\{v_k^s\}_{k=1}^m$ be the set of control points after s subdivision steps. To each control point v_k^s corresponds a basis function B_k^s whose support is made up of a set of $d+1$ consecutive knot intervals. We call the word made up of the labels of these intervals the *domain decomposition* c_k^s corresponding to the control point v_k^s. For brevity, we will often omit the adjective "domain", when no confusion is possible. The decompositions of two successive control points are *consistent*: the final d symbols of one word are identical to the initial d symbols of the following word.

A subdivided control point v_l^{s+1} has a non-zero coefficient in the mask of v_k^s if the support of B_l^{s+1} is contained in the support of B_k^s. Therefore the decomposition c_l^{s+1} of v_l^{s+1} is a *factor* (consecutive subsequence) of c_k^s rewritten using the rules of the grammar. In addition, the positions of the knots in the knot sequence of v_l^{s+1} can also be deduced from c_k^s up to a translation, and therefore the influence of v_k^s in the computation of v_l^{s+1} can be computed using the classical knot-insertion algorithm based on the blossom of each polynomial piece of the B-spline proposed by Ramshaw (1989). This is illustrated in the following subsection.

2.2.1. Example: Fibonacci system

We illustrate this computation of masks with the following L-system as an interval subdivision descriptor. This system was first studied by Leonardo Fibonacci (Levine and Steinhardt, 1986) and will therefore be referred to as the Fibonacci system. Because it creates tilings with two kinds of intervals, one long and one short, it is convenient to use the interval labels $\{L, S\}$ instead of $\{A_1, A_2\}$.

- Symbols: $\{L, S\}$.
- Rules: $\{L \to SL, S \to L\}$.

Associated interval lengths and ratio, satisfying (3), are

- Lengths: φ and 1 where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.
- Ratio: $\rho = \varphi$.

Note that the specific values of the lengths and of the ratio are determined by the choice of L-system, as will be discussed in detail in the next section. We take $d = 2$, and because the intervals are of positive length, it follows that the spline is C^1.

As an example, for a control point v_k^s of the decomposition LSL the rewritten decomposition is SLLSL, and the factors are SLL, LLS, LSL. Let us call the corresponding control points v_l^{s+1}, v_{l+1}^{s+1} and v_{l+2}^{s+1}. In terms of knot sequences, the knot sequence corresponding to the decomposition LSL is

$$t_k^s, \qquad t_{k+1}^s, \qquad t_{k+2}^s, \qquad t_{k+3}^s.$$

The subdivided knot sequence is

$$t_l^{s+1}, \qquad t_{l+1}^{s+1}, \qquad t_{l+2}^{s+1}, \qquad t_{l+3}^{s+1}, \qquad t_{l+4}^{s+1}, \qquad t_{l+5}^{s+1}$$

with

$$t_l^{s+1} = t_k^s, \qquad t_{l+2}^{s+1} = t_{k+1}^s, \qquad t_{l+3}^{s+1} = t_{k+2}^s, \qquad t_{l+5}^{s+1} = t_{k+3}^s.$$

From knot-insertion and the blossom formulation, each new control point can be written as a convex combination of old control points (Goldman, 1990). More precisely, each control point is given by the blossom p of one of the polynomial pieces it influences, evaluated at d consecutive knots:

$$v_l^{s+1} = p\big(t_{l+1}^{s+1}, t_{l+2}^{s+1}\big) = p\big(t_{l+1}^{s+1}, t_{k+1}^s\big)$$

and, using the multiaffinity of the blossom,

$$v_l^{s+1} = \frac{t_{k+2}^s - t_{l+1}^{s+1}}{t_{k+2}^s - t_k^s} p\big(t_k^s, t_{k+1}^s\big) + \frac{t_{l+1}^{s+1} - t_k^s}{t_{k+2}^s - t_k^s} p\big(t_{k+1}^s, t_{k+2}^s\big).$$

Using the decomposition information, we have

$$\begin{aligned} t_{k+2}^s - t_{l+1}^{s+1} &= 2\varphi, \\ t_{l+1}^{s+1} - t_k^s &= 1, \\ t_{k+2}^s - t_k^s &= 2\varphi + 1. \end{aligned}$$

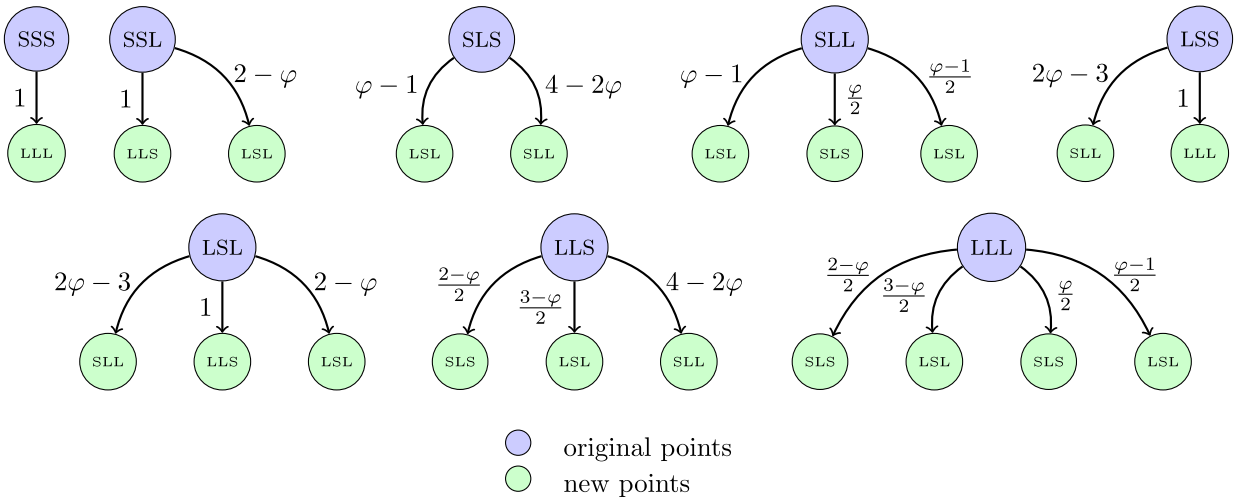

Fig. 1. Subdivision masks for the C^1 spline ($d = 2$) based on the Fibonacci system.

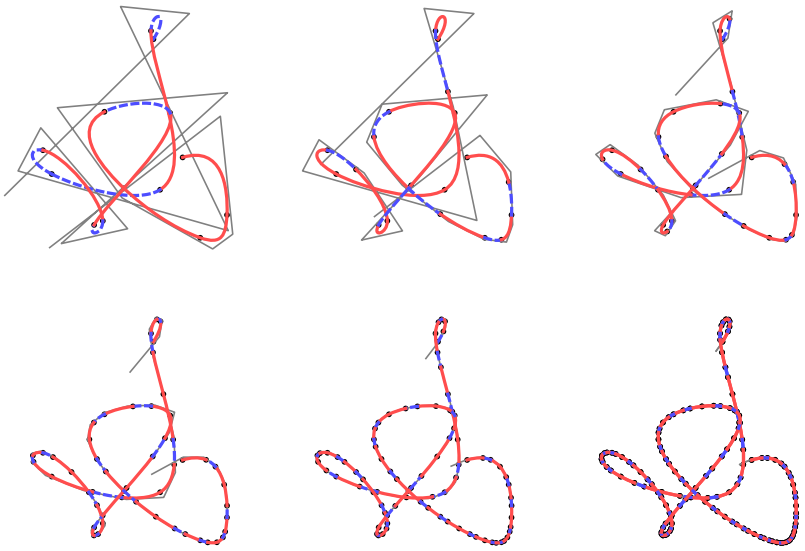

Fig. 2. Subdivision of a B-spline with C^2 continuity ($d = 3$) using the Fibonacci subdivision rules. The blue dashed curve segments are of parametric length S and the red solid ones L.

Thus, since $\varphi^2 = \varphi + 1$, we obtain

$$v_l^{s+1} = (4 - 2\varphi)v_{k-1}^s + (2\varphi - 3)v_k^s.$$

Similarly, we have

$$v_{l+1}^{s+1} = v_k^s,$$
$$v_{l+2}^{s+1} = (2 - \varphi)v_k^s + (\varphi - 1)v_{k+1}^s.$$

We can see that all the control points with an identical decomposition subdivide in the same way, and one mask can be computed per decomposition. Since the same ratio ρ scales down the interval lengths at every subdivision step, these masks are the same whatever the subdivision step.

Using this technique for all the decompositions, we obtain the full set of masks shown in Fig. 1. This technique generalises for any choice of d. For example, Fig. 2 illustrates the C^2 Fibonacci scheme with $d = 3$.

2.2.2. Reducing the number of masks to be computed

The number of decompositions and therefore the size of the full set of masks is n^{d+1} when n is the number of symbols and d the degree of the spline.

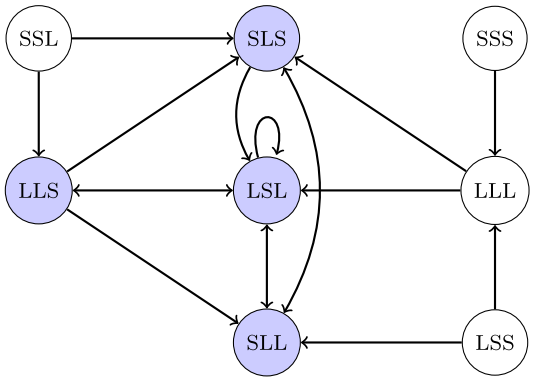

Fig. 3. Decomposition graph for the Fibonacci subdivision rules with $d = 2$ and continuity C^1. After a small number of subdivision steps, the blue-coloured decompositions are the only ones that will be used.

This size can however be reduced if we consider only the decompositions that can still appear after a few subdivision steps. Consider the graph whose vertices are the decompositions, with an oriented edge between two decompositions if the second one appears in the mask of the first one (Fig. 3). After a finite number of iterations, the only useful masks correspond to decompositions which belong to strongly connected components of this graph or which are reachable from such components. These last decompositions are the only factors of length $d + 1$ in the interval sequence after a few steps. In the case of the Fibonacci scheme, it can be demonstrated that their number is $d + 2$. Thus, the number of masks which are still relevant after a few steps is linear in the degree of the spline, rather than exponential. An important reference, for these and related questions, is Lothaire (2002).

In practice this decomposition graph will be useful if we compute the masks on the fly from a given axiom: only the masks corresponding to decompositions which belong to a strongly connected component of the graph, or which are reachable from such a component, need be stored for future use.

2.2.3. Stencils

In uniform subdivision, the subdivision mask collects, for a given control point, the coefficients of its contributions to the subdivided control points. In the opposite direction, the subdivision stencil collects for a given subdivided control point the set of coefficients weighting the control points used to compute it. (Note that some authors, especially in cases where there is no need to refer to what we have called a mask, write "mask" where we have used "stencil". Our terminology is fairly standard amongst authors who need both masks and stencils.)

We have seen in the example above how using an L-system as an interval subdivision descriptor allows us to generalise the notion of mask to non-uniform subdivision while defining one mask per domain decomposition. The notion of stencil, however, cannot be generalised solely using the domain decomposition. Indeed, for a given subdivided knot decomposition, the layout of the parents and their weights is not unique, so a unique stencil cannot be computed per decomposition. An example of a point with non-unique stencil is the one with decomposition LSL shown in Fig. 1.

2.3. Practical implementation

The user provides as input the degree d of the spline, and an L-system containing the set of intervals to be used in the subdivision scheme, along with their subdivision rules. The L-system is checked for validity: valid lengths $\{a_i\}$ and the ratio ρ are automatically computed as described in Section 3.

An initial control polygon must be specified, where each vertex in the control polygon has an associated word made up of $d + 1$ symbols. Consecutive words must be consistent, as described in Section 2.2, in order to be interpreted as domain decompositions. From these initial decompositions and the values of the interval lengths and the ratio, a small number of subdivision masks can be computed and stored.

When it comes to applying the subdivision mask, for each control point we must locate the subdivided control points it influences, and add its contributions using the weights of its mask. This can be done in different ways: details of one straightforward implementation are given in a worksheet for Sage mathematical software. The link http://www.sagenb.org/home/pub/3520 provides direct access to the commented code. In particular, boundaries are processed in a simple way, without any multiple knots: control points which are present in the mask of some existing control points, but with an insufficient number of parents to fully compute their value, are not computed.

3. Validity of an L-system as an interval subdivision descriptor

As seen in Section 2.1.2, for each rule of the form (1) for which A_i can actually occur, the lengths of labelled intervals and the ratio ρ must satisfy (3). Moreover, in order to show convergence, below, we require $\rho > 1$. Not every set of rules

leads to such lengths and ratio. For this reason we introduce the concept of a valid L-system, we give sufficient conditions on the rules for the L-system to be valid, and finally we discuss uniqueness of the lengths and the ratio for a valid L-system.

3.1. Validity

It is convenient to define the matrix $\mathbf{M} = [m_{ij}]_{n\times n}$ where m_{ij} is the number of A_j in the right side of the rule describing the splitting of A_i, and the vector of lengths $\mathbf{a} = [a_i]_{n\times 1}$. Eq. (3) can then be written as

$$\rho\mathbf{a} = \mathbf{M}\mathbf{a}. \tag{4}$$

In other words, an appropriate vector of lengths $\mathbf{a}$ and a ratio $\rho > 1$ should be, respectively, a real and positive eigenvector and an associated real eigenvalue of the matrix $\mathbf{M}$. This yields the following definition for a valid interval subdivision descriptor.

Definition 3.1 *(Valid L-system).* An L-system is a valid interval subdivision descriptor if the matrix $\mathbf{M}$ has at least one real eigenvalue $\rho > 1$ and an associated real and positive eigenvector $\mathbf{a}$.

The lengths and ratio do not depend on the degree d of the spline used for the definition of a subdivision scheme based on the L-system. Also, we emphasise that a positive eigenvector has all components strictly positive.

As an example of a valid L-system, consider the Fibonacci system introduced in Section 2.2.1:

- Symbols: $\{L, S\}$.
- Rules: $\{L \to SL, S \to L\}$.

The corresponding matrix is

$$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$$

and its eigen-elements $(\rho, \mathbf{a})$ are

$$\left\{\left(\varphi, \begin{bmatrix} \varphi \\ 1 \end{bmatrix}\right), \left(\psi, \begin{bmatrix} \psi \\ 1 \end{bmatrix}\right)\right\}.$$

The eigenvalue $\psi = \frac{1-\sqrt{5}}{2} < 0$ is not appropriate, but the golden ratio $\varphi = \frac{1+\sqrt{5}}{2} > 1$, which is the ratio used in Section 2.2.1 along with the entries of its associated eigenvector as interval lengths, provides a valid interval subdivision descriptor.

As an example of an invalid L-system, consider the following:

- Symbols: $\{A_1, A_2\}$.
- Rules: $\{A_1 \to A_1A_1A_2, A_2 \to A_2\}$.

The corresponding matrix $\mathbf{M}$ defined by the rules is

$$\mathbf{M} = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

and its eigen-elements $(\rho, \mathbf{a})$ are

$$\left\{\left(2, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right), \left(1, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)\right\}.$$

This system produces, after s subdivision steps, knot interval lengths equal to 2^{-s} or to 0, and may in fact lead to a reasonable subdivision scheme, but this question is outside the scope of this paper. As mentioned below, in the conclusion, relaxation of the positivity constraint will be studied in future work.

Theorem 3.2. *If the L-system provided as input is valid, then the sequence of polygonal lines defined by the subdivision scheme described in Section* 2.3 *converges uniformly to the B-spline curve.*

Proof. Since $\mathbf{a}$ and ρ are real and positive, the vector entries can be associated with interval lengths corresponding to symbols of the L-system, while the eigenvalue corresponds to the ratio between lengths in two successive steps, in the construction of the degree-d B-spline subdivision scheme described in Section 2.

We denote by $\{t_k^s\}_{k=1}^{m^s+d+1}$ the knot sequence after s subdivision steps, m^s being the number of control vertices after these subdivisions. The particular treatment of boundaries described in Section 2.3 shrinks the parameter domain but in such a

way that t^s_{d+1} and $t^s_{m^s+1}$ remain constant with respect to s. Note that $[t^s_{d+1}, t^s_{m^s+1}]$ is the interval on which the spline $\mathcal{S}$, which is preserved through the subdivision process, is properly defined. Assume that a_i is the largest component of $\mathbf{a}$. It follows that the longest possible interval length at this subdivision step is a_i/ρ^s. Since $\rho > 1$, we are sure that after a finite number of subdivision steps, the number of knots increases. Moreover, since the intervals are of positive length, the second derivative $\mathcal{S}''$ exists and is piecewise continuous for $d > 1$.

Then, from Theorem 3.3 in Cohen and Schumaker (1985), there exists an integer $S \geqslant 0$ such that

$$\left\|\mathcal{L}^s - \mathcal{S}\right\| \leqslant C\left(\frac{a_i}{\rho^s}\right)^2 \left\|\mathcal{S}''\right\|, \quad \text{for all } s \geqslant S,$$

where $\mathcal{L}^s$ is a piecewise linear parametrisation of the control polygon after s subdivision steps, C is a constant, and $\|.\|$ denotes the uniform norm. This completes the proof. □

3.2. Sufficient conditions for validity

The Perron–Frobenius theorem allows us to give useful sufficient conditions on the graph of the matrix $\mathbf{M}$ in order for the L-system to be valid. The following definitions, lemma and theorem can be found in Varga (1962).

Definition 3.3 *(Irreducible matrix).* For $n \geqslant 2$, an $n \times n$ matrix $\mathbf{M}$ is reducible if there exists an $n \times n$ permutation matrix $\mathbf{P}$ such that

$$\mathbf{PMP}^T = \begin{bmatrix} \mathbf{M}_{1,1} & \mathbf{M}_{1,2} \\ \mathbf{0} & \mathbf{M}_{2,2} \end{bmatrix},$$

where $\mathbf{M}_{1,1}$ is an $r \times r$ submatrix and $\mathbf{M}_{2,2}$ is an $(n-r) \times (n-r)$ submatrix, where $1 \leqslant r < n$. If no such permutation matrix exists, then $\mathbf{M}$ is irreducible.

If $\mathbf{M}$ is a 1×1 matrix, then $\mathbf{M}$ is irreducible if its single entry is non-zero, and reducible otherwise.

Definition 3.4 *(Strongly connected graph).* The graph $\mathcal{G}(\mathbf{M})$ of $\mathbf{M} = [m_{ij}]_{n\times n}$ is defined to be the directed graph on n nodes $\{N_1, N_2, \ldots, N_n\}$ in which there is a directed edge leading from N_i to N_j if and only if $m_{ij} \neq 0$.

The graph $\mathcal{G}(\mathbf{M})$ is called strongly connected if for each pair of nodes (N_i, N_k) there is a sequence of directed edges leading from N_i to N_k.

Lemma 3.5. *The matrix $\mathbf{M}$ is irreducible if and only if $\mathcal{G}(\mathbf{M})$ is strongly connected.*

Theorem 3.6 *(Perron–Frobenius). If $\mathbf{M} = [m_{ij}]_{n\times n}$ with $m_{ij} \geqslant 0$, is irreducible, then*

- *the spectral radius $\rho(\mathbf{M})$ is an eigenvalue of $\mathbf{M}$ and $\rho(\mathbf{M}) > 0$;*
- *there exists a positive eigenvector associated with $\rho(\mathbf{M})$;*
- *$\rho(\mathbf{M})$ increases when any value of $\mathbf{M}$ increases;*
- *$\rho(\mathbf{M})$ is a simple eigenvalue of $\mathbf{M}$.*

Further, the *Collatz–Wielandt formula*, which is true for all non-negative matrices, states that $\rho(\mathbf{M}) = \max_{\mathbf{x}\in\mathcal{N}} f(\mathbf{x})$ where

$$f(\mathbf{x}) = \min_{i\in\{1\ldots n\}|x_i\neq 0} \frac{[\mathbf{Mx}]_i}{x_i}$$

and $\mathcal{N} = \{\mathbf{x} \mid \mathbf{x} \geqslant \mathbf{0} \text{ with } \mathbf{x} \neq \mathbf{0}\}$.

A sufficient condition for an L-system to be valid is then given by the following theorem. Recall that the outdegree of a node in a directed graph is equal to the number of directed edges, including possibly a loop back to the node itself, going out from the node.

Theorem 3.7. *If the graph $\mathcal{G}(\mathbf{M})$ of the matrix $\mathbf{M} = [m_{ij}]_{n\times n}$ of the L-system is strongly connected and if there exists $i \in \{1, \ldots, n\}$ such that N_i has outdegree at least 2, then the L-system is valid for defining an interval subdivision descriptor.*

Proof. The graph $\mathcal{G}(\mathbf{M})$ is strongly connected, so the matrix $\mathbf{M}$ is irreducible. From the Perron–Frobenius theorem, the spectral radius $\rho(\mathbf{M})$ is an eigenvalue and there exists a positive eigenvector $\mathbf{a} = [a_i]_{n\times 1}$ associated with it. Let us show that the spectral radius satisfies $\rho(\mathbf{M}) > 1$.

Let i be such that the outdegree of N_i is at least 2. It follows that

$$\sum_{j=1}^{n} m_{ij} > 1.$$

We define the matrix $\mathbf{N} = [n_{ij}]_{n\times n}$ as

$$n_{ij} = \begin{cases} 1, & \text{if } m_{ij} \neq 0; \\ 0, & \text{otherwise.} \end{cases}$$

For any integer $s \in \{1, \dots, n\}$, let $\mathbf{N}^s = [n_{ij}^{(s)}]_{n\times n}$ be the matrix $\mathbf{N}$ multiplied s times by itself, and similarly for $\mathbf{M}^s$. Then $n_{ij}^{(s)}$ is equal to the number of paths in the graph $\mathcal{G}(\mathbf{M})$ leading from the node N_i to the node N_j with s directed edges. Since m_{ij} is a non-negative integer we have $m_{ij}^{(s)} \geqslant n_{ij}^{(s)}$.

Since the graph $\mathcal{G}(\mathbf{M})$ is strongly connected, there is a sequence of directed edges leading from any node N_l to the node N_i. Let s be the number of edges which constitute this path for a given l, and consider only values of s such that $s < n$. We have

$$\begin{aligned}
&n_{li}^{(s)} \geqslant 1, \\
&m_{li}^{(s)} \geqslant 1, \\
&\sum_{k=1}^{n} m_{lk}^{(s)} m_{kj} \geqslant m_{ij}, \\
&\sum_{j=1}^{n}\sum_{k=1}^{n} m_{lk}^{(s)} m_{kj} \geqslant \sum_{j=1}^{n} m_{ij}, \\
&\sum_{j=1}^{n} m_{lj}^{(s+1)} > 1.
\end{aligned}$$

Since all rows of $\mathbf{M}$ contain at least one non-zero value (otherwise it would be reducible) the sum of elements of any row of $\mathbf{M}^s$ does not decrease with s. Therefore, for all $l \in \{1, \dots, n\}$,

$$\sum_{j=1}^{n} m_{lj}^{(n)} > 1.$$

Since $\mathbf{M}^n$ is non-negative, the Collatz–Wielandt formula states that

$$\rho\left(\mathbf{M}^n\right) \geqslant f(\mathbf{1})$$

with

$$f(\mathbf{x}) = \min_{i\in\{1\dots n\}|x_i\neq 0} \frac{[\mathbf{M}\mathbf{x}]_i}{x_i}$$

and $\mathbf{1}$ the vector with all n entries equal to 1. Therefore,

$$\left(\rho(\mathbf{M})\right)^n \geqslant \rho\left(\mathbf{M}^n\right) \geqslant \min_{l\in\{1\dots n\}} \sum_{j=1}^{n} m_{lj}^{(n)} > 1, \tag{5}$$

and $\rho(\mathbf{M}) > 1$. As a consequence, the eigen-elements $(\rho(\mathbf{M}), \mathbf{a})$ establish the validity of the L-system. □

A useful corollary may be added to this theorem.

Corollary 3.8. *Suppose $n \geqslant 2$. If the graph $\mathcal{G}(\mathbf{M})$ of the matrix $\mathbf{M} = [m_{ij}]_{n\times n}$ of the L-system is strongly connected and if, in at least one rule, the symbol on the left-hand side is also in the word on the right-hand side, then the L-system is valid for defining an interval subdivision descriptor.*

This corollary can be applied, for example, to the Fibonacci system, with the symbol L in the rule $L \to SL$.

Proof. Let $i \in \{1, \dots, n\}$ be such that A_i is a symbol which is in both sides of one of the rules of the L-system. Then a directed edge loops from the node N_i to itself in the graph $\mathcal{G}(\mathbf{M})$. Furthermore this graph is strongly connected, and $n \geqslant 2$, so there is at least one edge which goes from N_i to another node, which leads to the result. □

In order to show that strong connectedness of $\mathcal{G}(\mathbf{M})$ is not alone sufficient for the L-system to be valid, we consider the following L-system.

- Symbols: $\{A_1, A_2\}$.
- Rules: $\{A_1 \to A_2, A_2 \to A_1\}$.

The corresponding matrix is

$$\mathbf{M} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and its eigen-elements $(\rho, \mathbf{a})$ are

$$\left\{ \left(1, \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right), \left(-1, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) \right\}.$$

The graph $\mathcal{G}(\mathbf{M})$ is strongly connected, but we do not have $\rho(\mathbf{M}) > 1$.

Next, note that the conditions in Theorem 3.7 are not necessary since the following L-system is valid even though its graph is not strongly connected.

- Symbols: $\{A_1, A_2\}$.
- Rules: $\{A_1 \to A_1A_2, A_2 \to A_2A_2A_2\}$.

The corresponding matrix is

$$\mathbf{M} = \begin{bmatrix} 1 & 1 \\ 0 & 3 \end{bmatrix}$$

and its eigen-elements $(\rho, \mathbf{a})$ are

$$\left\{ \left(1, \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right), \left(3, \begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) \right\}.$$

The second pair of eigen-elements makes the L-system valid.

3.3. Properties of eigen-elements of a valid L-system

We now establish certain uniqueness results, and give an interpretation of the left eigenvector of $\mathbf{M}$.

3.3.1. Uniqueness of the eigenvalue

We show that even without the hypothesis of strong connectedness of $\mathcal{G}(\mathbf{M})$, we can make strong statements about the uniqueness of an eigenvalue that establishes validity.

Theorem 3.9. *If $(\rho, \mathbf{a})$ establishes the validity of the L-system, then ρ is unique and equal to the spectral radius $\rho(\mathbf{M})$.*

Proof. The matrix $\mathbf{M}$ is non-negative, so it can be obtained as the limit of positive matrices. It therefore follows from the Perron–Frobenius theorem that $\mathbf{M}$ has an eigenvalue equal to its spectral radius, and an associated eigenvector with non-negative components. The transpose $\mathbf{M}^T$ is non-negative because $\mathbf{M}$ is non-negative, and the spectral radii of the two matrices are the same. We can therefore apply the previous statement to the transpose of $\mathbf{M}$: let $\mathbf{s}$ be a non-negative eigenvector of $\mathbf{M}^T$, corresponding to the eigenvalue $\rho(\mathbf{M}) = \rho(\mathbf{M}^T)$.

We have

$$\mathbf{s}^T\mathbf{M} = \rho(\mathbf{M})\mathbf{s}^T.$$

Since $\rho > 1$ is a real eigenvalue of $\mathbf{M}$ with a positive eigenvector $\mathbf{a}$, we have

$$\rho(\mathbf{M})\mathbf{s}^T\mathbf{a} = \mathbf{s}^T\mathbf{M}\mathbf{a} = \mathbf{s}^T\rho\,\mathbf{a}.$$

The vector $\mathbf{s}$ is not zero, and $\mathbf{s}$ is non-negative, so it has at least one positive component, and $\mathbf{a}$ is positive. Consequently, $\mathbf{s}^T\mathbf{a} > 0$ and $\rho(\mathbf{M}) = \rho > 1$. □

3.3.2. The right eigenvector

A simple sufficient condition for uniqueness of the eigenvector is that the eigenvalue $\rho(\mathbf{M})$ be a simple root of the characteristic equation. In particular, from the Perron–Frobenius theorem, if the graph corresponding to the system is strongly connected, then the eigenvector is unique.

However, the eigenvector $\mathbf{a}$ for a valid system is not necessarily unique, if the graph corresponding to the system is not strongly connected. For example, consider the L-system defined by

- Symbols: $\{A_1, A_2\}$.
- Rules: $\{A_1 \to A_1A_1, A_2 \to A_2A_2\}$.

The corresponding matrix is

$$\mathbf{M} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

and any positive vector **a** is an eigenvector corresponding to the double eigenvalue $\rho(\mathbf{M}) = 2$. This non-uniqueness is also illustrated by the non-uniform subdivision schemes proposed by Sederberg et al. (1998), which have $2\mathbf{I}$ as matrix, where **I** is the identity matrix with as many rows as edges in the initial control polygon. The only eigenvalue is $\rho(\mathbf{M}) = 2$ but the elements of the eigenvector **a** are any knot-interval lengths provided by the user.

3.3.3. The left eigenvector

Finally, let us consider the left eigenvector of the matrix **M** associated with ρ:

$$\rho\mathbf{s}^T = \mathbf{s}^T\mathbf{M}. \tag{6}$$

The component s_j gives the steady-state value for the number of symbols A_j used at each step, and consequently the steady-state value for the length resulting from the use of the symbol A_j, for each j. In fact, suppose that at a particular step of the subdivision process, the symbol A_i appears n_i times. At the next step, for each j, these n_i symbols will generate intervals corresponding to A_j that have length $n_i m_{ij} a_j/\rho$. The total length (resulting from symbols of all types i) associated with symbols A_j at the next step is therefore equal to $\sum_{i=1}^{m} n_i m_{ij} a_j/\rho$, and in steady state this must be equal to $n_j a_j$. Thus, the n_i define the components of the left eigenvector **s**.

It is unlikely that particular methods and their associated axioms will lead to steady-state behaviour in the sense just described, especially (but not only) because the n_i must be integral. Consider for example the Fibonacci system. The left eigenvector of **M** associated with the spectral radius φ is $\mathbf{s}^T = \begin{bmatrix} 1/\varphi & 1/\varphi^2 \end{bmatrix}$, and so the system does not exhibit steady-state behaviour for any choice of axiom, i.e., for any initial choice of symbols L and S.

Although steady-state behaviour is unlikely, the vector **s** appears in the proof of Theorem 3.9 given above, and it therefore seems worthwhile to have an interpretation of this vector.

3.4. Example

In this section we illustrate these results, related to the validity of L-systems as interval subdivision descriptors, by means of a system which permits a ratio ρ as close to 1 as desired. Such a system can be used to create a subdivision which grows as slowly as desired.

A ratio of this kind can be achieved by introducing a delay in the splitting of most intervals: only one kind of interval splits at a time, while the others wait in a delay queue. The following L-system, with $n \geqslant 2$, creates such a low-growth subdivision.

- Symbols: $\{A_1, A_2, \ldots, A_n\}$.
- Rules: $\{A_1 \to A_nA_n, A_i \to A_{i-1} \text{ for all } i > 1\}$.

The corresponding matrix is

$$\mathbf{M} = \begin{bmatrix} 0 & \cdots & 0 & 2 \\ 1 & \ddots & \ddots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

The graph $\mathcal{G}(\mathbf{M})$ is strongly connected and $\sum_{j=1}^{n} m_{1j} = 2 > 1$, so this L-system is valid as an interval subdivision descriptor. Its spectral radius is $\rho(\mathbf{M}) = 2^{\frac{1}{n}}$, where the associated eigenvector is $[a_i]_{n\times 1}$ with $a_i = (\rho(\mathbf{M}))^{n-i}$. By increasing n, the ratio can be set as close to 1 as desired.

The example just presented illustrates an extreme case: it was chosen for its transparency. It is clearly possible to introduce a wide variety of strategies with ratios as close to 1 as desired. Such strategies will be discussed in a later paper.

4. Conclusion

We have described how an L-system can be used as an interval subdivision descriptor if its matrix of rules has an eigenvalue strictly greater than 1 and an associated positive eigenvector. We have provided sufficient but not necessary conditions on the graph of the matrix of rules for the L-system to be valid. Furthermore, if this interval subdivision descriptor describes

the subdivision of a knot sequence of a non-uniform B-spline of a given degree d, then the well-known knot-insertion rules produce a non-uniform subdivision scheme that we have shown to be uniformly convergent. We have also shown that this subdivision scheme is defined by a finite set of masks (at most n^{d+1} where n is the number of symbols in the alphabet of the L-system). Finally we have provided an implementation of such schemes which is available as a worksheet for Sage software.

In future work we will study more carefully necessary conditions on the L-system for the subdivision scheme to converge. In particular it would be of interest to relax the constraint on the positivity of interval length, and on the uniform ratio for all interval subdivisions. These generalisations would be useful, for example, in the context of the specification of high-level operations that permit adaptive control of the subdivision process. Furthermore, such non-uniform subdivision schemes generalise to tensor-product subdivision surfaces. Handling extraordinary vertices, however, will require us to develop analyses and implementations specific to that case. We will also study the speed of convergence of the non-uniform schemes discussed relative to classical schemes. Finally, the possibility of having subdivision schemes with a finite set of rules, a ratio lower than 2, and irregular or pseudo-periodic sampling properties, will be studied in the context of multi-resolution analysis.

References

Barry, P.J., Zhu, R.F., 1992. Another knot insertion algorithm for B-spline curves. Computer Aided Geometric Design 9, 175–183.

Boehm, W., 1980. Inserting new knots into B-spline curves. Computer-Aided Design 12, 199–201.

Cashman, T.J., Dodgson, N.A., Sabin, M.A., 2009a. Selective knot insertion for symmetric, non-uniform refine and smooth B-spline subdivision. Computer Aided Geometric Design 26, 472–479.

Cashman, T.J., Dodgson, N.A., Sabin, M.A., 2009b. A symmetric, non-uniform, refine and smooth subdivision algorithm for general degree B-splines. Computer Aided Geometric Design 26, 94–104.

Cohen, E., Lyche, T., Riesenfeld, R., 1980. Discrete B-splines and subdivision techniques in computer-aided geometric design and computer graphics. Computer Graphics and Image Processing 14, 87–111.

Cohen, E., Schumaker, L.L., 1985. Rates of convergence of control polygons. Computer Aided Geometric Design 2, 229–235.

Goldman, R.N., 1990. Blossoming and knot insertion algorithms for B-spline curves. Computer Aided Geometric Design 7, 69–81.

Goldman, R., Warren, J.D., 1993. An extension of Chaiken's algorithm to B-spline curves with knots in geometric progression. CVGIP: Graphical Model and Image Processing 55, 58–62.

Herman, G.T., Lindenmayer, A., Rozenberg, G., 1974. Description of developmental languages using recurrence systems. Theory of Computing Systems 8, 316–341.

Lane, J.M., Riesenfeld, R.F., 1980. A theoretical development for the computer generation and display of piecewise polynomial surfaces. IEEE Trans. Pattern Analysis Machine Intell. 2, 35–46.

Levine, D., Steinhardt, P.J., 1986. Quasicrystals. I. Definition and structure. Phys. Rev. B 34, 596–616.

Lothaire, M., 2002. Algebraic Combinatorics on Words. Cambridge University Press.

Prusinkiewicz, P., Hanan, J., 1989. Lindenmayer Systems, Fractals, and Plants. Lecture Notes in Biomathematics, vol. 79. Springer-Verlag, Berlin.

Prusinkiewicz, P., Samavati, F., Smith, C., Karwowski, R., 2002. L-system description of subdivision curves. International Journal of Shape Modeling 9, 41–59.

Ramshaw, L., 1989. Blossoms are polar forms. Computer Aided Geometric Design 6, 323–358.

Schaefer, S., Goldman, R., 2009. Non-uniform subdivision for B-splines of arbitrary degree. Computer Aided Geometric Design 26, 75–81.

Sederberg, T.W., Zheng, J., Sewell, D., Sabin, M., 1998. Non-uniform recursive subdivision surfaces. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques. ACM, New York, NY, USA, pp. 387–394.

Varga, R.S., 1962. Matrix Iterative Analysis. Prentice-Hall.

Velho, L., 2003. Stellar subdivision grammars. In: SGP'03: Proceedings of the 2003 Eurographics/ACM SIGGRAPH symposium on Geometry processing, pp. 188–199.